



The Future of Affirmative Action: Reclaiming the Innovative Deal

Citation

Susan Sturm & Lani Guinier, The Future of Affirmative Action: Reclaiming the Innovative Deal, 84 Calif. L. Rev. 953 (1996).

Published Version

<http://scholarship.law.berkeley.edu/californialawreview/vol84/iss4/3/>

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:12967841>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

July 1996

The Future of Affirmative Action: Reclaiming the Innovative Deal

Susan Sturm

Lani Guinier

Follow this and additional works at: <http://scholarship.law.berkeley.edu/californialawreview>

Recommended Citation

Susan Sturm and Lani Guinier, *The Future of Affirmative Action: Reclaiming the Innovative Deal*, 84 CAL. L. REV. 953 (1996).
Available at: <http://scholarship.law.berkeley.edu/californialawreview/vol84/iss4/3>

This Article is brought to you for free and open access by the California Law Review at Berkeley Law Scholarship Repository. It has been accepted for inclusion in California Law Review by an authorized administrator of Berkeley Law Scholarship Repository. For more information, please contact jcera@law.berkeley.edu.

The Future Of Affirmative Action: Reclaiming The Innovative Ideal

Susan Sturm†
Lani Guinier‡

We are witnessing a broad-based assault on affirmative action—in the courts,¹ the legislatures,² and the media.³ Opponents have defined affirmative action as a program of racial preferences that threatens fundamental American values of fairness, equality, and democratic opportunity.⁴ Opponents successfully depict racial preferences as

Copyright © 1996 Susan Sturm and Lani Guinier.

† Professor of Law, University of Pennsylvania.

‡ Professor of Law, University of Pennsylvania.

We deeply appreciate the helpful comments and suggestions of many colleagues on earlier drafts of this article, including Mary Frances Berry, James Crouse, Dayna Cunningham, Harlon Dalton, George Madaus, Paul Minorini, Martha Minow, William Eric Perkins, Michael Selmi, and participants in the Diversity of Language Seminar at the University of Pennsylvania and the New Paradigms Symposium, supported by the Ford Foundation and the Annenberg Public Policy Center. We are also grateful for the research assistance of Marc Kesselman and Monifa Williams.

1. See *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469 (1989) (plurality opinion); *Hopwood v. Texas*, 78 F.3d 932 (5th Cir. 1996) (holding that a law school may not use diversity as a basis for taking race into account in law school admissions), *cert. denied*, 116 S. Ct. 2582 (1996); *Podberesky v. Kirwan*, 38 F.3d 147 (4th Cir. 1994) (striking down the use of race-based scholarships), *cert. denied*, 115 S. Ct. 2001 (1995).

2. See Steven A. Holmes, *As Affirmative Action Ebbs, A Sense of Uncertainty Rises*, N.Y. TIMES, July 6, 1995, at A1 (discussing former Senate majority leader and presidential candidate Bob Dole's drafting of legislation intended to "end special considerations of race in Federal programs"); Donna St. George, *For White Men, Anger Taking Political Shape*, PHILA. INQUIRER, Nov. 12, 1995, at A1 (describing efforts in the Illinois, Georgia, and Pennsylvania state legislatures to end race and gender preferences). In the words of Constance Horner, a member of the U.S. Civil Rights Commission, "It's the end of an era. What tells us that this is the end of an era is that all branches of government—the courts, the Congress, the White House, even the state legislatures—are actively engaged in the same process, and that's a rare event in American politics." *Id.*

3. See Janine Jackson, *White Man's Burden: How the Press Frames Affirmative Action*, EXTRA!, Sept.-Oct. 1995, at 7 (describing press tendency to define affirmative action using the "hot-button" term "quota," to ignore realities of continuing discrimination against women and people of color, and to portray questions of discrimination "as most importantly the concern of white men"); Robert M. Entman, *Accentuating the Negative: Media Coverage of Affirmative Action* (Nov. 15, 1995) (unpublished manuscript, on file with authors).

4. See, e.g., HERMAN BELZ, *EQUALITY TRANSFORMED: A QUARTER-CENTURY OF AFFIRMATIVE ACTION* 17-18 (1991) (arguing that racial preferences distort the color-blind ideal originally contemplated by the sponsors of the Civil Rights Act of 1964); Morris B. Abram, *Affirmative Action: Fair Shakers and Social Engineers*, 99 HARV. L. REV. 1312 (1986) (arguing that color-conscious remedies exacerbate racial tensions and contravene the original goal of the civil rights

extraordinary, special, and deviant—a departure from prevailing modes of selection. They also proceed on the assumption that, except for racial or gender preferences, the process of selection for employment or educational opportunity is fair, meritocratic, and functional.⁵ Thus, they have positioned affirmative action as unnecessary, unfair, and even un-American.⁶

Those of us pursuing the quest of racial and gender justice in a genuinely democratic society face a crucial challenge. How do we respond to this assault on affirmative action? How do we invite a deeper conversation and analysis of selection and admissions conventions in pursuit of fairness? Understandably, much of the response has been reactive. Supporters of affirmative action typically engage the debate on the terms defined by the assault: affirmative action must continue. It is fair. It is still needed to rectify continued exclusion and marginalization in the society.⁷

Supporters of affirmative action have also put forward a critique of the fairness and functionality of existing merit standards.⁸ They mar-

movement to promote an equal-opportunity society); Terry Eastland, *The Case Against Affirmative Action*, 34 WM. & MARY L. REV. 33 (1992) (calling for a return to race-neutral principles).

5. See, e.g., BRON R. TAYLOR, *AFFIRMATIVE ACTION AT WORK: LAW, POLITICS, AND ETHICS* (1991) (observing from extensive personal interviews that staunch opponents of affirmative action typically view affirmative action as departing from the merit principle); Eastland, *supra* note 4, at 40 (criticizing affirmative action for departing from "objective" merit); Allan C. Ornstein, *Quality, Not Quotas*, SOCIETY, Jan./Feb. 1976, at 10, 17 ("Using quotas instead of quality to select people for jobs and promotions rewards the dumb, lazy, and unambitious at the expense of the smart, talented, and ambitious.").

6. See, e.g., Clint Bolick, *Discriminating Liberals*, N.Y. TIMES, May 6, 1996, at A15; Nicholas Lemann, *Taking Affirmative Action Apart*, N.Y. TIMES, June 11, 1995, §6 (Magazine), at 36, 40 (discussing how opponents of affirmative action, after honing their arguments for 30 years, now argue that affirmative action threatens basic American values).

7. See Roy L. Brooks, *The Affirmative Action Issue: Law, Policy, and Morality*, 22 CONN. L. REV. 323, 359-62 (1990) (arguing that affirmative action "helps to make American society a more just society"); Luke C. Harris & Uma Narayan, *Affirmative Action and the Myth of Preferential Treatment: A Transformative Critique of the Terms of the Affirmative Action Debate*, 11 HARV. BLACKLETTER J. 1, 4 (1994) (arguing that affirmative action seeks to counter some of the continuing effects of a historical experience that impedes equal opportunity today and "to promote fairness, equality, and full citizenship by affording members of excluded groups a fair chance to enter significant social institutions"); Benjamin L. Hooks, *Affirmative Action: A Needed Remedy*, 21 GA. L. REV. 1043, 1043 (1987) ("Until a nondiscriminatory society exists, affirmative action will remain a needed remedy. . ."); Alex M. Johnson, Jr., *Defending the Use of Quotas in Affirmative Action: Attacking Racism in the Nineties*, 1992 U. ILL. L. REV. 1043, 1073 ("Considering the black experience in a historical and contextual framework that focuses not only on the past, but [also on] the future of American society, the use of quotas is the most efficacious method for achieving racial equality in contemporary American society.").

8. See, e.g., IRIS M. YOUNG, *JUSTICE AND THE POLITICS OF DIFFERENCE* 204 (1990) (arguing that merit criteria are "normative and cultural rather than neutrally scientific"); Sheila Foster, *Difference and Equality: A Critical Assessment of the Concept of "Diversity"*, 1993 WIS. L. REV. 105, 157 ("Recognizing that institutional standards and norms were developed in the context of the power struggles that have historically existed between certain groups in this society calls into question the assumption that they are neutral or universally constructed."); Robert L. Hayman, Jr., *Re-Cognizing*

shal considerable evidence showing that these standards exclude women and people of color, and that people who were excluded in the past do not yet operate on a level playing field.⁹ They have also challenged the justification for relying on these exclusionary criteria; they argue that the selection criteria do not predict the future performance of candidates in the positions they seek to occupy.¹⁰ They then rely on this critique of the fairness and validity of existing merit standards to justify departing from those standards for women and people of color. Affirmative action is justified to level the playing field, to rectify the biases built into the existing selection system, and to remedy past and continuing exclusion or underrepresentation.¹¹

Despite the moral and empirical force of these arguments, there is a sense in which they are not being heard. They certainly have not reshaped the terms or tone of the public debate. The most compelling moral claims are simply dismissed as special-interest pleading. Part of the reason for this asymmetry is that proponents of racial and gender justice have responded to the debate only as it is framed by the current assault.

This narrow response has tactical, strategic, and substantive costs. As a tactical matter, proponents have accepted a paradigm that misdirects attention and energy into trench warfare, rather than into pursuing a progressive agenda. By reacting defensively to the current onslaught, they have foreclosed discussion of new, innovative strategies for racial and gender justice. Substantively, they have accepted an existing

Inequality: Rebellion, Redemption and the Struggle for Transcendence in the Equal Protection of the Law, 27 HARV. C.R.-C.L. L. REV. 9, 48-49 (1991) (arguing that appeals to "race-neutral" concepts such as "merit" and "color-blindness" tacitly adopt current structures of racial privilege); Mark Kelman, *Concepts of Discrimination in "General Ability" Job Testing*, 104 HARV. L. REV. 1158, 1159 (1991) (concluding that the use of tests is inconsistent with meritocratic principles that mainstream conservatives voice in interpreting the antidiscrimination norm); Duncan Kennedy, *A Cultural Pluralist Case for Affirmative Action in Legal Academia*, 1990 DUKE L.J. 705 (challenging notion that "merit" is neutral, impersonal, and somehow developed outside the forces of social power and its attendant categories of race, gender, and class); Michael Selmi, *Testing for Equality: Merit, Efficiency, and the Affirmative Action Debate*, 42 UCLA L. REV. 1251, 1276-77 (1995) (challenging central assumption of affirmative action critics that test scores are closely correlated with productivity and that small test score differentials indicate meaningful performance differentials); see also Gary Peller, *Race Consciousness*, 1990 DUKE L.J. 758, 779 (explaining that standards of "merit" have been constructed from particular perspectives and as responses to specific historical circumstances).

9. See *infra* notes 169-187 and accompanying text.

10. See *infra* Part II.A.

11. See *infra* notes 48-52, 196-200, 218-219 and accompanying text; see also Lemann, *supra* note 6, at 43:

The history of affirmative action can be seen as a struggle over the fairness of the modern meritocracy, with minorities arguing that educational measures shouldn't be the deciding factor in who gets ahead and opponents of affirmative action saying that to bend the criteria for blacks is to discriminate unfairly against more deserving whites.

framework of selection that is fundamentally and deeply flawed for those whom it includes as well as for those left out.

In other words, affirmative action, as it is currently practiced, supplements an underlying framework of selection that is implicitly arbitrary and exclusionary. It does not challenge the overall operation of a conventional and static selection process; instead, it creates exceptions to that process. Those exceptions play into existing racial stereotypes, predictably generating backlash.¹² By implicitly legitimizing a selection process that operates in the name of merit, affirmative action programs reinforce that backlash. Programs perceived as racial preferences also enable employers to cast issues of economic retrenchment in terms of racial conflict. Many white workers who acknowledge the lack of corporate responsibility for the economic well-being of workers still focus their wrath and blame on the workers perceived as beneficiaries of affirmative action.¹³

It is time, we argue, for those of us committed to racial and gender equity to advance a more fundamental critique of existing selection and admission conventions. It is time to discuss how conventional assessment and predictive criteria do not function fairly, democratically, or even meritocratically for many Americans who are not members of racial or gender minorities. To reclaim the moral high ground, we must broaden and expand the terms of engagement. By revealing faulty assumptions about the concept of affirmative action and the system of selection in which it operates, we can move from an incrementalist strategy of inclusion for a few to a transformative vision of reform for the many.

To reopen the conversation on race, gender, and democratic opportunity, it is necessary to change the paradigm. Certainly, we must challenge out loud the basic assumption that affirmative action is a departure from an otherwise sound meritocracy. At the same time, we must challenge existing add-on practices of affirmative action as too conservative a remedy. The experience of women and people of color

12. The extent of the backlash, however, may be exaggerated. See Lini S. Kadaba, *Study Finds No Affirmative Action Backlash*, PHILA. INQUIRER, Dec. 12, 1995, at A2 (discussing research by a Pennsylvania State University professor indicating that "whites who have seen [affirmative action] in action hold less prejudiced views than others"); Entman, *supra* note 3, at 1, 21-22 (contending that the media's focus on conflict overstates the extent of public opposition to affirmative action).

13. St. George, *supra* note 2 (quoting a Ukrainian immigrant's son, a police officer, and a Chicago lawmaker who says Illinois government "has used affirmative action to create a new old-boys' network. They're just different boys."). Despite all the changes in the economy, "politicians, bosses, coworkers blam[e] affirmative action as the reason whites can't get ahead." *Id.* at A17; Louis Uchitelle, *Union Goal of Equality Fails the Test of Time*, N.Y. TIMES, July 9, 1995, at A1, A18 (discussing a 41-year-old white male ironworker who resents affirmative action, despite being "welcomed with open arms into this union" as the son and grandson of ironworkers and "[having] had it pretty easy" in the job market).

offers insights beyond showing how and why those particular people have been excluded. We need to show that the current one-size-fits-all ranking system of predicting "merit" is no longer justified or productive for anyone.

The present system of selection is unfair for people who are neither women nor people of color. It denies opportunity for advancement to many poor and working-class Americans of all colors and genders who could otherwise obtain educational competence. It is underinclusive of those who can actually do the job. It is deeply problematic as a predictor of actual job performance. Across-the-board, it does violence to fundamental principles of equity and "functional merit"¹⁴ in its distribution of opportunities for admission to higher education, entry-level hiring, and job promotion.

Typical among the existing criteria and selection methods are paper-and-pencil tests, such as the Scholastic Assessment Test (SAT), the Law School Admissions Test (LSAT), and civil service exams.¹⁵ These tests, which are used to predict future performance based on existing capacity or ability, do not correlate with future performance for most applicants, at least not as a method of ranking those "most qualified."¹⁶ These tests and informal criteria making up our "meritocracy" tell us more about past opportunity than about future accomplishments on the job or in the classroom.¹⁷

In challenging the way these tests are used, we are not proposing a critique of merit per se. Nor are we advancing an entirely original argument. Simply stated, we seek to highlight the way that certain paper-and-pencil tests have been used as "wealth preferences" or poll taxes to determine who gets to participate as full citizens in our democracy.¹⁸ As Michael Lind argues in a slightly different context, these tests are used, in conjunction with subjective assessments and informal networks, to develop a class-linked opportunity structure that credentializes "a social oligarchy."¹⁹

The approach we develop in this Article links affirmative action initiatives with the project of fundamentally rethinking how we define and practice genuine merit selection. We argue that affirmative action is

14. The term "functional merit" normally refers to the qualities needed to complete a job or perform it competently. The term, as we use it, is further defined later. See *infra* notes 36-37, 53-54, and accompanying text.

15. See *infra* Part II.A for a discussion of these tests.

16. See *infra* Part II.A.

17. See *infra* notes 148-168 and accompanying text (discussing the link between parental socioeconomic status and test performance).

18. See *infra* Part IV.E (explaining the connection between conventional merit selection and barriers to universal citizenship).

19. Michael Lind, *Prescriptions for a New National Democracy*, 110 *POL. SCI. Q.* 563, 582 (1995-96).

an opportunity to take from the margin to rethink the whole. Affirmative action is not about exceptions to the norm; it is about the norm itself. Affirmative action, and the experience of those who have been previously excluded, provide a window on a much larger set of questions. These are the same questions that companies and educational institutions must face to meet the demands of an economy in transition: Can we define and predict ability to perform based on one-size-fits-all tests and criteria? How do we go about identifying the type of worker/student who will perform successfully under changing economic conditions? Is sameness fairness? Or must we reconsider the notion that in a complicated world there are simple and single solutions? How do we rethink the process and content of selection to better accommodate the demands of the twenty-first-century workplace?

It is time to ask a different set of questions about affirmative action, questions that address the most pressing problems facing not only people of color and women, but all of those who are unfairly excluded from participation in work and education. We need to go beyond the modest curative of affirmative action to examine more deeply our system of selecting and evaluating all workers and students. This approach to affirmative action can open up an inquiry into the adequacy and legitimacy of the one-size-fits-all approach to selection that prevails in many arenas. In this way, affirmative action provides a less reactive, more transformative critique that highlights the range of preferences implicit in conventional selection and prediction criteria.

Patterns of exclusion experienced primarily by women and people of color are, nevertheless, still important. They serve as signals. Patterns of race- and gender-based exclusion signal the possibility that bias or unfair advantage has operated in the ostensibly neutral selection process. They also signal the inadequacy of traditional methods of selection for everyone, and the need to rethink the process used to allocate opportunities to participate in work and school. In other words, patterns of exclusion provide a window on the methods for "inclusion." They are an important source of continuous critique of monolithic and monochromatic ranking and selection processes.

Rethinking our assumptions about selection is important to be able to pursue goals of racial and gender justice and fairness. Even more, it is crucial to our capacity to develop productive, fair, and efficient institutions that can meet the challenges of a rapidly changing, unstable, and increasingly complex marketplace. By using the experience of those on the margin to rethink the whole, we may forge a new, progressive vision of cross-racial collaboration, functional diversity, and genuinely democratic opportunity.

Part I of this Article sets out the stock narratives underlying the affirmative action debate and the assumptions about merit and fairness that underlie those narratives. Part II then shows that the current meritocracy is neither fair nor functional, and that it in effect gives preference to candidates who enjoy privileged socio-economic positions. Part III articulates the need for a new paradigm for recruitment, selection, and promotion. The goals of this new framework are threefold: 1) to locate and develop workers who can do the job, 2) to attain genuine inclusion of underrepresented groups, and 3) to promote a collaborative opportunity structure that brings fresh perspectives to doing a better job. Part IV describes one alternative: a framework for selection that shifts the focus from prediction to experience, based on structured, participatory, and accountable assessment.²⁰ In this final Part, we suggest that new paradigms for affirmative action can integrate diversity and merit, and thereby build into the framework of selection the capacity to adapt to the innovative challenges of a dynamic and uncertain economy.

I

THE STOCK AFFIRMATIVE ACTION NARRATIVES

In this Section we make visible the premises and assumptions that typically frame the affirmative action debate. Many employers and educational institutions select applicants by ranking them through a combination of paper-and-pencil tests and subjective assessments.²¹ De-

20. Other possible approaches include the use of institution-specific measures to set a floor of standards, above which schools and companies would select a diverse group of people who meet the minimum qualifications. These ideas are further described *infra* at Part IV.A.

Similar efforts to create collaborative and innovative alternatives influenced our Fall 1995 seminar, *Critical Perspectives on the Law: Issues of Race and Gender*, in which students actively engaged in the process of unpacking and reconstituting the language and framework of affirmative action. To discuss ways of transforming the debate over affirmative action, the seminar used the example of Lowell High School, a magnet public high school in San Francisco. This elite school became "a battleground in the war over racial preferences" when a Chinese-American group challenged a desegregation consent decree, which required the school to maintain a racial and ethnic balance. Elaine Woo, *Caught on the Wrong Side of the Line?: Chinese Americans Must Outscore All Other Groups to Enter Elite Lowell High in San Francisco, Sparking an Ugly Battle Over Diversity and the Image of a 'Model Minority'*, L.A. TIMES, July 13, 1995, at A1.

Our seminar used a variety of approaches to create a dynamic framework for open, engaged, and constructive dialogue about race, gender, and the relationship of affirmative action to legal problem-solving. We designed the class to explore substantive issues, but also to provide students with a structured opportunity to communicate across differences and collaborate in building innovative approaches to rethinking race and gender issues. Participants wrote weekly reflection pieces on the material assigned, often in response to a question or problem posed by student facilitators of that particular session. With the participants' permission, we quote in this Article from particular reflection pieces that exemplify the creative and sophisticated work done by students in that seminar. Several of these reflection pieces are cited below.

21. See HOWARD GARDNER, *MULTIPLE INTELLIGENCES: THE THEORY IN PRACTICE* 163-64 (1993) (citation omitted):

cision makers equate performance on these "neutral" selection criteria with merit. The assumption is that this system is fair and functional, that applicants have an unbiased opportunity to compete for positions, and that the resultant meritocracy operates to identify the people who are most qualified for and deserving of the position.

We begin by sketching out the prototypical selection processes at issue in the affirmative action debate. We describe how affirmative action typically operates within these scenarios. We also identify the particular conceptions of fairness and merit that permeate the debate, and show that they are both overly narrow and historically recent. This discussion lays the foundation for the subsequent demonstration that these underlying premises about selection are both unfair, in that they arbitrarily exclude some people and advantage others, and invalid, in that they fail to define either the goals or attributes of successful performance or to predict in most cases the individuals who can meet them.

A. *The Narratives*

Competing narratives drive the affirmative action debate. Each story is propelled by different assumptions about fairness and merit.²² Each story proceeds from different assumptions about the baseline of decision making: how fair, unbiased, and merit-driven is the system in which affirmative action operates? Although many of these differences in assumptions are never directly expressed, they implicitly provide the analytical framework for the argument advanced by the story. The stock story of affirmative action critics in the employment context (and the one that appears most often in the cases) is of the white civil servant—say a police officer or firefighter—John Doe. He scored several points higher on the civil service exam and interview rating process, but lost out to a woman or person of color who did not score as high on those selection criteria.²³ John Doe claims, along with many public op-

[T]here has been a virtual mania for producing tests for every possible social purpose. In addition to standardized tests for students, we have such tests for teachers, supervisors, soldiers, and police officers; we use adaptations of these instruments to assess capacities not only in standard areas of the curriculum but also in civics and the arts; and we can draw on short-answer measures for assessing personality, degrees of authoritarianism, and compatibility for dating.

22. Robert S. Chang, *Reverse Racism!: Affirmative Action, the Family, and the Dream that Is America*, 23 HASTINGS CONST. L.Q. (forthcoming 1996) (manuscript at 5, on file with authors) (noting that different sides in the affirmative action debate have vastly different notions of fairness and merit that inform their construction of affirmative action).

23. See, e.g., *Johnson v. Transp. Agency*, 480 U.S. 616 (1987); *Wygant v. Jackson Bd. of Educ.*, 476 U.S. 267 (1986). Many of the affirmative action cases under Title VII involve public employees such as police, firefighters, and teachers. The most politicized version of the anti-affirmative action narrative is typified by the campaign strategy used by Senator Jesse Helms, the white incumbent, against Harvey Gant, the black challenger: "The Helms campaign commercial displayed a white working class man tearing up a rejection letter while the voice-over said, 'You

ponents of affirmative action, that he is more qualified for the job, and that it is unfair to allow race or gender considerations to deprive him of what he "deserves."²⁴

The recent decision of the United States Court of Appeals for the Fifth Circuit, *Hopwood v. Texas*, exemplifies the stock affirmative action narrative in the education context.²⁵ The University of Texas Law School based its admissions decisions largely on an applicant's score on the Texas Index ("TI"), a composite of undergraduate grade point average ("GPA") and Law School Aptitude Test ("LSAT") score. The Law School used this number to rank candidates for admission, to predict their performance in the first year of law school, and to determine the number of offers needed to fill its class.²⁶ To increase the efficiency of the admissions process, the law school sorted the applicant pool into three categories according to applicants' TI scores: presumptive admit, presumptive deny, and a discretionary zone. The law school supplemented this system with an affirmative action program designed to remedy past discrimination in the Texas school system and to increase the diversity of the law school. The law school's affirmative action pro-

needed that job, and you were the best qualified. . . . But it had to go to a minority because of a racial quota." Cheryl I. Harris, *Whiteness as Property*, 106 HARV. L. REV. 1707, 1767 n.261 (1993) (quoting ANDREW HACKER, *TWO NATIONS: BLACK AND WHITE, SEPARATE, HOSTILE, UNEQUAL* 202 (1992)); see also Chang, *supra* note 22 (manuscript at 7).

24. See St. George, *supra* note 2, which tells the story of Brian Gilhooly, a white Chicago firefighter, who took an exam for promotion to lieutenant:

Gilhooly had studied for months. He'd gotten his associate degree in fire science. He was sure he'd make it.

The ruddy firefighter ranked 175th on a list of 2,059. But when rankings were adjusted for race, Gilhooly's number fell to 217—which nudged him out of the running.

"It was like, they can't do this to me," he recalls. "This can't be." His father was a policeman—and made rank, three times. The promotion to fire lieutenant was then worth \$6,459 a year in salary, and it was a step toward higher promotions.

. . . .

He recalls a scene at the firehouse, when his daughter, Maeve, then 6, was visiting. The young girl eyed a black paramedic and whispered: "Daddy, is that the man who took your job?"

"She didn't know—all she saw was black skin," Gilhooly says. "It was a sobering moment." . . .

. . . He's doing better than a lot of people in America, though not better than his father—and not what he feels he deserves. He can't get over the idea that test scores aren't utmost.

Id. at A17; see also Louis Harris, *Affirmative Action and the Voter*, N.Y. TIMES, July 31, 1995, at A13 ("Most white[s surveyed] thought that when minority members are given preferential treatment in hiring, an unqualified black is given a job that a qualified white man deserves—reverse discrimination.").

25. *Hopwood v. Texas*, 78 F.3d 932 (5th Cir. 1996), *cert. denied*, 116 S. Ct. 2582 (1996).

26. *Id.* at 935. The law school also considers the qualities each applicant might bring to the class, which could include an applicant's background, life experiences, and outlook. *Id.*

gram used lower ranges to place black and Mexican Americans into the three admissions categories.²⁷

Cheryl Hopwood, a white applicant who was rejected by the law school, scored higher on the TI than some black and Mexican-American applicants who were admitted.²⁸ She, like John Doe, claimed that she was more qualified for admission to the Law School, and that it was unfair to deny her what she "deserved."

Both of these narratives depict the type of affirmative action programs that have come to define the debate: outcome-oriented programs that establish numerical goals to increase the participation of underrepresented groups in various settings.²⁹ These programs establish their goals by determining the percentage of group members (i.e., women, blacks, Latinos, etc.) in the pool from which candidates are drawn. In the educational and employment context, race and gender generally operate as a plus factor in the selection process.³⁰ These programs are frequently referred to, by both supporters and opponents of affirmative action, as racial preferences.³¹ Thus, the stock story frames the affirmative action debate in terms of racial preferences that depart from normal, universal, unbiased, and purportedly fair standards for determining merit.

27. *Id.* at 935-36. The presumptive admit score in 1992 for whites was 199, while for African and Mexican Americans it was 189. The presumptive deny score for whites was 192, while for minorities it was 179. *Id.* at 936.

28. *Id.* at 937-38. She also scored higher than over 100 white applicants who were admitted, but this fact is conveniently dropped from the stock narrative. *State of Texas Petition for Certiorari, Texas v. Hopwood*, 116 S. Ct. 2582 (1996).

29. Affirmative action debates involve programs in such diverse settings as schools, the workplace, government contracting, and voting. *See* Holmes, *supra* note 2 (citing affirmative action legal cases and controversies involving a university scholarship program, a school desegregation plan, a ballot initiative to halt the use of racial preferences in hiring and college admissions, the awarding of federal contracts, and the drawing of legislative districts with a majority of minority voters).

30. *See, e.g., Johnson v. Transportation Agency*, 480 U.S. 616 (1987) (upholding affirmative action plan for hiring and promotion that used race and gender as plus factors); *Regents of the Univ. of Cal. v. Bakke*, 438 U.S. 265 (1978) (holding that race is a valid consideration in admissions process).

31. *Harris & Narayan, supra* note 7, at 13; *see also* GERTRUDE EZORSKY, *RACISM AND JUSTICE: THE CASE FOR AFFIRMATIVE ACTION* (1991) (focusing her discussion of affirmative action on "[n]umerical goals for hiring and promoting minorities which permit preference to 'basically qualified' blacks"); *Harris, supra* note 24 (finding that preferential treatment, among people Harris surveyed, connoted "favoritism or nepotism"); David Hollinger, *Group Preferences, Cultural Diversity, and Social Democracy: Notes Toward a Theory of Affirmative Action*, 55 REPRESENTATIONS 31 (1996).

*B. Merit and Fairness in the Stock Affirmative
Action Narratives*

In essence, John Doe and Cheryl Hopwood are advancing the same two claims: 1) they have more merit than beneficiaries of affirmative action, and 2) as a matter of fairness, they are entitled to the position for which they applied. These two claims rest on the premise that qualifications should determine allocation of employment and educational opportunity, that existing selection criteria determine who is most qualified, and that departures from conventional merit standards are unfair.³²

This argument implicitly defines merit and fairness in particular ways. John Doe and Cheryl Hopwood assume that they "merit" the coveted position.³³ One possible basis for this claim rests on the notion of desert: because they were next in line, based on established criteria of selection, they deserve the position.³⁴ They may also base their claim on the idea of earned recognition: "when an individual has worked hard and succeeded (by her own or others' measures), [she] deserves recognition, praise and/or reward."³⁵ The claim of merit in the stock narrative could also proceed from a functional idea of merit: someone who has the qualities needed to perform effectively in the position under consideration.³⁶ Many affirmative action critics equate functional

32. See Stanley Fish, *Reverse Racism or How the Pot Got to Call the Kettle Black*, ATLANTIC MONTHLY, Nov. 1993, at 128.

I can hear the objection in advance: "What's the difference? Unfair is unfair: you [a white male] didn't get the job; you didn't even get on the short list." . . .

... [T]he unfairness scenarios are presented as simple contrasts between two decontextualized persons who emerge from nowhere to contend for a job or a place in a freshman class. Here is student A; he has a board score of 1,300. And here is student B; her board score is only 1,200, yet she is admitted and A is rejected. Is that fair? Given the minimal information provided, the answer is of course no.

Id. at 136.

33. The narrative assumes that jobs and educational opportunities should be allocated based on merit. See Richard H. Fallon, Jr., *To Each According to His Ability, From None According to His Race: The Concept of Merit in the Law of Antidiscrimination*, 60 B.U. L. REV. 815 (1980).

34. See Selmi, *supra* note 8, at 1252-53 (noting that employment decisions are often based on small point score differentials without consideration as to whether the differences are meaningful); Guy Watts, *Affirmative Action: Reframing the Discourse 1* (Dec. 4, 1995) (unpublished reflection piece, on file with the authors) ("The problem with the debate is that the only issue considered is merit based on grades and test scores. The two concepts have become almost synonymous.").

35. Laura K. Bass, *Affirmative Action: Reframing the Discourse 1* (Dec. 4, 1995) (unpublished reflection piece, on file with the authors).

36. See Nicole Perkins, *Affirmative Action: Reframing the Discourse 1* (Dec. 4, 1995) (unpublished reflection piece, on file with the authors):

My definition of qualified starts from the bottom up as opposed to the top down. In other words, I see qualified as meaning the minimum capabilities necessary to get the job done. If you can get the job done, you're qualified to do the job. Anything else you can (or cannot) do is irrelevant. For example, if the job requires adding two plus two, then anyone who can add two plus two is qualified. Qualified does not mean that the cut off should be at

merit with a numerical ranking on standard paper-and-pencil tests. Those with the higher scores are presumably those who can best perform in the position under consideration.³⁷ John Doe and Cheryl Hopwood could also assert that they are better in some abstract or universalistic sense—that the selection process measures a quality that society generally deems valuable.³⁸ In conventional terms, those with easily testable or quantifiable types of intelligence are more highly valued and will contribute to the overall quality of the institution, and thus are more qualified.³⁹

The stock narrative's claim of unfairness builds on these assumptions that merit should and, in the absence of affirmative action, does govern employment and educational decision making. To the extent that affirmative action departs from an otherwise fair and valid system of selection, it is unfair. Fairness, like merit, is also a concept with varying definitions.⁴⁰ The stock story defines fairness in terms of formal or procedural fairness to the individual. This view of fairness, which we call fairness-as-sameness, emphasizes the importance of treating everyone the same, giving everyone the same formal opportunity to enter the competition for a position, and evaluating each person's results the same way. If everyone takes the same test, and every applicant's test is evaluated in the same manner, then the test is fair. Because affirmative action evaluates some people's test results differently, it is unfair.

This notion of fairness also implicitly involves a concept of notice and detrimental reliance. It is only fair that the rules governing selection remain constant throughout the process. If an employer or school conveys the expectation that it will select candidates based on test results and scores, and it then departs from those standards to take race or gen-

the top percentage of applicants who can do a number of things up and above the task of adding two plus two.

37. The criteria tested could include analytical ability, general aptitude, speed and strength, or knowledge of the particular job. See, e.g., *Guardians Ass'n v. Civil Serv. Comm'n*, 630 F.2d 79 (2d Cir. 1980), *cert. denied*, 452 U.S. 940 (1981).

38. Richard Fallon has offered a similar typology of merit: "[A] person's merit changes not only in a relative but in an absolute sense as the complementary utility of his attributes [vary] across time and place" and, it would seem, institutional context. Fallon, *supra* note 33, at 828.

39. Along with these differing conceptions of merit, the debate puts forward differing goals for the process of determining merit. One could evaluate a merit system based on its capacity to reward the most deserving candidates, to create incentives for future applicants to develop certain skills and abilities, to eliminate arbitrariness and bias and promote the appearance of fairness, to reduce costs of selection, to help define and promote the organization's mission, or to serve more general societal goals.

40. For a discussion of different definitions of fairness in the context of testing and of the implications of those different definitions, see COMMITTEE ON THE GENERAL APTITUDE TEST BATTERY, NATIONAL RESEARCH COUNCIL, *FAIRNESS IN EMPLOYMENT TESTING: VALIDITY GENERALIZATION, MINORITY ISSUES, AND THE GENERAL APTITUDE TEST BATTERY* 253-60 (John A. Hartigan & Alexandra K. Wigdor eds., 1989) [hereinafter *FAIRNESS IN EMPLOYMENT TESTING*].

der into account, it is breaching an implicit agreement to use particular criteria for hiring or admissions.⁴¹ A crucial premise of this fairness challenge to affirmative action is the assumption that tests afford equal opportunity to demonstrate individual merit, and therefore are not biased. The presumption is that one-size-fits-all selection produces fair results.

The yardstick is the implicit, and sometimes explicit, metaphor for the version of merit and fairness that underlies the stock narrative.⁴² The assumption is that institutions know what they are looking for (height), they know how to measure who has those characteristics (yards, meters), they can fairly replicate the measurement process (using the ruler), and they can rank people accordingly (by height). This approach to merit selection has been institutionalized in a manner that rests heavily on the use of standardized tests and other "objective" screening tools. Those who participate in selection, along with those who challenge how affirmative action operates, assume that these approaches to selection are generally valid and fair, and that they are justified as a means of pursuing merit. They assume that objective tests for particular attributes of merit, perhaps supplemented by subjective methods such as unstructured interviews and reference checks, are the state-of-the-art selection method, that they can be justified as predictive of performance, that they are the most efficient method of selection, and that no better alternative exists.

The stock narrative naturalizes the current yardstick model of merit. Thus, "more qualified" means that the applicant scored higher on standardized tests and interview ratings. Fairness means that applicants should be numerically ranked for selection using "objective," race- and gender-neutral selection criteria.

In fact, the dominance of standardized tests in selection is a relatively recent development. The civil rights revolution, and the introduction of affirmative action programs, occurred at the same time that society was formalizing a "meritocracy" based on education and standardized testing.⁴³ The construction of this meritocracy was part of an

41. Cf. Samuel Issacharoff, *When Substance Mandates Procedure: Martin v. Wilks and the Rights of Vested Incumbents in Civil Rights Consent Decrees*, 77 CORNELL L. REV. 189, 193, 222-23 (1992) (arguing that employees' expectations with respect to seniority and other "objective criteria" constitute a property interest deserving constitutional protection); Samuel Issacharoff, *Reconstructing Employment*, 104 HARV. L. REV. 607, 621-24 (1990) (book review) (discussing reliance-based interests of incumbent employees in their employment).

42. See, e.g., ROBERT G. CAMERON, *THE COMMON YARDSTICK: A CASE FOR THE SAT 2* (1989) ("The need arose for a common yardstick to measure school outcomes in college admissions, and the SAT has become that yardstick.").

43. JAMES CROUSE & DALE TRUSHEIM, *THE CASE AGAINST THE SAT* 31-37 (1988) (tracing the ascendance of standardized testing in the 1950's and linking it to criticism of progressivism and rising emphasis on efficiency and standardization); Leniann, *supra* note 6, at 42 (noting that, until the

overall rationalization and formalization of the selection process,⁴⁴ in part as a response to the simultaneous legalization of norms for workplace conduct.⁴⁵ This move to objective testing arguably minimized both arbitrariness and individualized bias. In fact, many defended objective testing on the grounds that it opened up opportunity to people of all socio-economic backgrounds.⁴⁶ It eliminated the class-linked prerequisites to work and education that governed in the pre-standardized testing era, and instituted a system that presumably offered everyone a fair, unbiased, and equivalent chance to compete for educational and employment opportunities.⁴⁷

Yet, for blacks and other people of color, "rationality" was introduced into an environment that was not benign. When the civil service and educational institutions moved to standardize admission criteria, most jobs and institutions of higher learning were still segregated.⁴⁸

1950s, standardized tests were uncommon and "there was no system in place that could evaluate and assign a numerical value to every American").

Although widespread reliance on testing for education and employment occurred in the 1950s, interest in standardized measurements of ability predated the post-war period. The use of formal testing has been traced to the work of Alfred Binet and his associates on intelligence testing at the turn of the century. GARDNER, *supra* note 21, at 163. Binet's work on developing standardized assessments for elementary school children led to the first intelligence tests, and became a dominant feature in the American educational and assessment landscape. *Id.* The nation's first mass intelligence testing occurred during World War I. CROUSE & TRUSHEIM, *supra*, at 21. For an eloquent and in-depth account of America's preoccupation with unitary measures of intelligence, see STEPHEN J. GOULD, *THE MISMEASURE OF MAN* (rev. ed. 1996).

44. CROUSE & TRUSHEIM, *supra* note 43, at 33 ("Schools would achieve new efficiency in the use of human resources by adopting the methods of the new science of psychometrics.").

45. For a discussion of this legalization process, see Lauren B. Edelman, *Legal Environments and Organizational Governance: The Expansion of Due Process in the American Workplace*, 95 AM. J. SOC. 1401, 1402, 1406-09 (1990) (arguing that the "civil rights movement and the mandates of the 1960s created a normative environment" that threatened the legitimacy of arbitrary organizational governance and precipitated the development of formal mechanisms of dispute resolution). However, the bureaucratization of the American workplace had been increasing for some time. See James N. Baron et al., *War and Peace: The Evolution of Modern Personnel Administration in U.S. Industry*, 92 AM. J. SOC. 350, 359-77 (1986) (arguing that unionization and increased government intervention between the 1920s and 1940s contributed to development of a more bureaucratic employment relationship).

46. CROUSE & TRUSHEIM, *supra* note 43, at 36 (noting that the ETS claimed that "[b]ecause the SAT is a uniform, color-blind test for predicting success, minority students and students from poor families would stand on equal footing with white middle-class students.").

47. *Id.* at 35-36:

"Unlike the personal interview, the classroom test, or the teacher's subjective evaluation, the objective test is a common touchstone. . . . It gives all students who take it the same chance, asks them to run the same race—even though they have had different economic backgrounds, different educational, cultural, and social opportunities."

Id. (quoting EDUCATIONAL TESTING SERVICE, 1960-61 ANNUAL REPORT 25-26 (1961)) (footnote omitted).

48. See *Brown v. Board of Education*, 347 U.S. 483 (1954); *Sweatt v. Painter*, 339 U.S. 629 (1950); ANDREW HACKER, *TWO NATIONS: BLACK AND WHITE, SEPARATE, HOSTILE, UNEQUAL* (1992). Indeed, the tests that were at issue in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), the case establishing that Title VII of the Civil Rights Act "proscribes not only overt discrimination but

Both in the educational context and the workplace, opportunities had been virtually closed to blacks, except for those institutions made up almost entirely of blacks. By itself, rationality would not cure the hostility and exclusion, which were justified by some as a "rational" response to black inferiority.⁴⁹

Thus, affirmative action as a remedial response emerged when many of society's norms were being challenged. On the one hand, merit was increasingly judged on a single or dominant criterion of performance—the ability to get good grades or perform well on tests that are designed to assess general intelligence or inherent ability.⁵⁰ On the other hand, that single criterion of performance was exactly the area in which blacks had been made most vulnerable, factually, legally, and mythologically. As a factual and legal matter, blacks' educational opportunities had been severely limited. As the subject of political and pseudo-scientific mythmakers, blacks were pervasively stereotyped as possessing less general intelligence or inherent ability.⁵¹ Thus, the push for greater rationality in the workplace and institutions of higher learning must be juxtaposed against efforts by blacks and other people of color to challenge both the racially discriminatory allocation of benefits and the racial stereotypes that reinforced that allocation of benefits.⁵²

also practices that are fair in form, but discriminatory in operation," were adopted the day after the enactment of the Act.

49. See *Southern Education Foundation Report Cites Test Misuse as a Cause of Segregation*, FAIRTEST EXAMINER, Fall/Winter 1995-96, at 7 ("A recent Southern Education Foundation (SEF) report cites the misuse of standardized admissions tests as one reason for the failure of desegregation in the region's state university systems."); cf. CROUSE & TRUSHEIM, *supra* note 43, at 19-21 (describing early role of intelligence testing in enabling universities to "retain their traditional clientele in the face of demographic changes"); Fish, *supra* note 32, at 135 (finding that the SAT tests were developed by "an out-and-out racist," Carl Campbell Brigham, "to confirm racist assumptions," and that Brigham had classified "American society into four distinct racial strains, with Nordic blue-eyed, blond people at the pinnacle and the American Negro at the bottom").

50. See CROUSE & TRUSHEIM, *supra* note 43, at 33-35 (documenting the rise of standardized tests as the benchmark of merit); GARDNER, *supra* note 21, at 163-65 (tracing the dominance of the uniform-testing model of performance back to the work on intelligence testing).

51. See, e.g., CHARLES MURRAY & RICHARD J. HERRNSTEIN, *THE BELL CURVE* (1994) (claiming that studies of similarly situated groups indicate that blacks, on average, have significantly lower intelligence or cognitive ability than whites). Such studies have been severely criticized. See, e.g., NEW REPUBLIC, Oct. 31, 1994, *passim*.

52. For example, educational institutions allocated benefits primarily to white males. Cf. Lemann, *supra* note 6, at 43 (noting that the development of a "numerical, education-based meritocracy was bad news for blacks [because] it apportioned opportunity on the basis of performance in the one area where blacks were most disadvantaged: education."). Indeed, prior to 1970, the University of Texas Law School admitted all applicants who took the LSAT and had at least a 2.0 or 2.2 grade point average. *Hopwood v. Texas*, 861 F. Supp. 551, 557 (W.D. Tex. 1994), *rev'd on other grounds*, 78 F.3d 932 (5th Cir.), *cert. denied*, 116 S. Ct. 2582 (1996).

Employers and labor unions allocated benefits primarily to working-class whites. See Uchitelle, *supra* note 13, at A18 (noting that, until after World War II, most unions were virtually closed to black workers and that as recently as 1964, the International Association of Machinists allowed its local chapters to exclude blacks).

Consequently, we will examine the story of merit and unfairness from several perspectives. In Part II.A, we scrutinize the assumption that conventional approaches to selection in fact distinguish those who can perform well from those who cannot. In Part II.B, we examine the fairness of the so-called meritocracy, showing that existing selection processes undervalue many important skills and capacities, give undue advantages to those from higher socio-economic backgrounds, and exclude women and people of color at a disproportionate rate.

II

UNPACKING MERIT, FAIRNESS, AND THE "TESTOCRACY"

The stock narrative proceeds from and depends on the premise that the selection criteria and processes used to rank applicants for jobs and admission to schools are basically fair and valid. Yet, a substantial body of literature fundamentally challenges this basic premise. Even accepting the definitions of merit and fairness employed in the stock narrative, current approaches to selection are extremely limited in their predictive capability. If we apply a more comprehensive conception of fairness, the existing "meritocracy" fails even more miserably as a method of selection.

We argue that the "meritocracy" is neither fair nor democratic, neither genuinely predictive nor functionally meritocratic. Not everyone is being given an equal opportunity to compete. Not everyone who could do the job, or who could bring new insights into how to do the job even better, is being given an opportunity to perform or succeed.

Instead, a "testocracy" masquerades as a meritocracy. By testocracy we refer to test-centered efforts to score applicants, rank them comparatively, and then predict their future performance. Although subjective and idiosyncratic measures are often used as part of the assessment process, such criteria frequently supplement the basic approach of seeking to predict future open-ended performance through static, closed-book, timed paper-and-pencil assessments of past ability. These approaches to selection are neither fair nor functional. The yardstick metaphor simply does not hold up under scrutiny.

A. Merit and the Fiction of Functionality

This Section scrutinizes the claim that the conventional "meritocracy" functions to identify those who can best perform in the positions under consideration. For purposes of our argument, we accept the idea, without question, that functional capacity to perform, or functional merit, is a legitimate consideration in distributing jobs and educa-

tional opportunities.⁵³ We also assess merit as a functional concept related to the capacity to perform effectively, rather than as a concept of desert or societal values. The concept of merit as desert or earned recognition collapses either into a concept of formal fairness or of functional merit.⁵⁴ The concept of merit as societal or institutional values is harder to defend against a claim of exclusion, unless it has a functional connection to the position under consideration. Merit as a functional value offers the strongest and most widely embraced justification for embracing standards that exclude members of particular disempowered groups.

The stock affirmative action narratives equate merit with performance on standardized tests. These paper-and-pencil tests do not fulfill their stated function. They do not reliably identify those applicants who will succeed in college or later in life, nor do they consistently predict those who are most likely to perform well in the jobs they will occupy. Used alone or in combination with informal networking and subjective assessment, timed paper-and-pencil tests screen out applicants who could nevertheless do the job.

To fulfill their stated function, testers must be able to identify and measure successful performance in the job or at school.⁵⁵ However, in both contexts, testers have failed to develop meaningful measures of

53. For arguments suggesting that consideration of merit in the distribution of opportunity will, in the current political context, simply reinforce existing unequal power relationships, see John Calmore, *Critical Race Theory, Archie Shepp, and Fire Music: Securing an Authentic Intellectual Life in a Multicultural World*, 65 S. CAL. L. REV. 2129, 2219 (1992) (linking merit standards to "tendencies, skills, or attributes of white America"); Richard Delgado, *Rodrigo's Chronicle*, 101 YALE L.J. 1357, 1364 (1992) ("Merit sounds like white people's affirmative action! . . . A way of keeping their own deficiencies neatly hidden while assuring that only people like them get in."); John E. Morrison, *Colorblindness, Individuality, and Merit: An Analysis of the Rhetoric Against Affirmative Action*, 79 IOWA L. REV. 313, 333 (1994) ("When Euro-Americans call the standards they meet neutral or objective, they validate themselves by showing that they would have succeeded, even if others did not have all the advantages they did."); Nancy S. Ehrenreich, *Pluralist Myths and Powerless Men: The Ideology of Reasonableness in Sexual Harassment Law*, 99 YALE L.J. 1177, 1234 (1990) ("[T]he prevailing ideology [based on meritocracy] systematically ignores differences among the citizenry as a whole, promoting a homogenous vision of American society that both excludes those groups who do not fit the accepted American model and elevates a small but powerful elite to the status of universal 'type.'"); Johnson, *supra* note 7, at 1070-71:

[The] belief in the concept of merit that is premised on the use of objective standards that allegedly can be fairly applied to discriminate between the deserving and the undeserving is not only factually inapposite, but is premised on a foundational claim of "acontextualism" that rejects the thoroughly contextually dependent nature of merit and standards.

54. The argument proceeds: I deserve this position because I was next in line. This is either an argument that the line arranges people in order of their capacity to perform (merit), or that it is simply unfair to depart from the stated rules of distribution (formal fairness). See also *infra* Part IV.E (discussing societal value of democracy as an additional basis for rethinking conventional ideas about merit).

55. See Kelman, *supra* note 8, at 1208 (1991) ("[W]hen test validators correlate a predictor (the screening test) with a criterion (performance on the job), the information they produce is useful only if the criterion they measure is a reasonable surrogate for actual productivity.").

successful performance. In the employment area, they have not attempted to correlate test performance with worker productivity or even with pay.⁵⁶ Instead, they rely on correlations between test performance and job sample tests or supervisor evaluations. Job test samples do not themselves correlate with productivity. Supervisor ratings are notoriously unreliable measures and have been shown to be biased in ways that correlate with race and gender.⁵⁷

In the educational context, testers attempt to correlate standardized tests with first-year performance in college or post-graduate education.⁵⁸ But this measure does not reflect successful overall academic achievement or performance in the areas valued by the educational institution.⁵⁹ Neither of the major testing services even attempts "to estimate the effect on college completion of different admissions policies, nor have they encouraged colleges to do so themselves."⁶⁰ Moreover, the data does not even consider the relationship between standardized test scores and a truly functional baseline—post-graduate "success" in life.⁶¹ Education and employment tests do not predict whether students will graduate or employees will succeed in their jobs.

Even if we accept the inadequate definitions of success used in conjunction with standardized tests, research shows a tenuous connection between test scores and successful performance. Recent studies show that the measured relationship between a test and predicted job performance, referred to as the correlation coefficient, is weak.⁶² The best employment tests have correlations of approximately .3.⁶³ Using the widely accepted statistical methodology for determining how much explanatory information a test provides,⁶⁴ a correlation of .3 means that "the test explains only 9% of the variation in predicted performance.

56. *Id.*

57. *See id.* at 1211; Selmi, *supra* note 8, at 1268.

58. *See* CROUSE & TRUSHEIM, *supra* note 43, at 40 (1988) ("Usually, high school GPA (or rank) is combined with a test score (or scores) to predict college freshman GPA.").

59. *See id.* at 186 n.23 ("Colleges do not use the SAT to forecast the academic or social value-added they produce, or their applicants' accomplishments in nonacademic endeavors."). Nor does the putative relationship between test scores and first-year grades take account of those who take the first two years of college to acclimate or to compensate for inadequate high school preparation.

60. *Id.* at 58.

61. Indeed, studies attempting such analyses have discovered that there is very little relationship between test performance and success in professional or public life. *See, e.g., infra* notes 92-104 and accompanying text.

62. For an excellent discussion of the correlation and predictiveness of tests, see Selmi, *supra* note 8, at 1262-70.

63. A recent assessment of the General Aptitude Test Battery found an average observed validity of about .22, with plausible adjustments for criterion unreliability raising the average observed validity to .25. FAIRNESS IN EMPLOYMENT TESTING, *supra* note 40, at 169.

64. "The square of the correlation coefficient can therefore be interpreted as the proportion of the total variation in the one variable explained by the other." Selmi, *supra* note 8, at 1263 n.37 (quoting HUBERT M. BLALOCK, JR., SOCIAL STATISTICS 409 (2d ed. 1972)).

In other words, the test leaves unexplained 91% of the variance reflected in the performance measure."⁶⁵

Validity studies of aptitude tests used to predict performance as measured only by first-year grades show correlations similar to those in the employment context. A recent study of the correlation of SAT scores with freshman grades showed correlations ranging from .32 to .36.⁶⁶ As David Owens notes, the correlation between "SAT scores and college grades . . . is lower than the correlation between weight and height; in other words, you would have a better chance of predicting a person's height by looking only at his weight than you would of predicting his freshman grades by looking only at his SAT scores."⁶⁷ A recent study of the University of Pennsylvania Law School found that LSAT scores were weak predictors of performance in law school. LSAT explained 21% of the differences in third-year grades. For first- and second-year students, it explained even less: 14% and 15% respectively.⁶⁸ A study of the Texas Index at issue in the *Hopwood* case also

65. *Id.* at 1263-64.

66. See WARREN W. WILLINGHAM ET AL., PREDICTING COLLEGE GRADES: AN ANALYSIS OF INSTITUTIONAL TRENDS OVER TWO DECADES 43 (1990).

67. DAVID OWEN, NONE OF THE ABOVE: BEHIND THE MYTH OF SCHOLASTIC APTITUDE 207 (1985); see also Fish, *supra* note 32, at 132 ("[W]hat is being measured by the SAT is not absolutes like native ability and merit but accidents like birth, social position, access to libraries, and the opportunity to take vacations or to take SAT prep courses.").

In March of 1994, the College Board introduced targeted modifications of the SAT, which included changing the name to the "Scholastic Assessment Test"; giving students about 15 seconds longer on each question; including longer reading and more comprehension passages; including 10 "grid-on" questions, which require students to figure out math problems without multiple choice answers; and eliminating the antonym section. Rachel Shteir, *Get Smart: Did the SAT-Makers Wise Up with a New, Unbiased Test?*, VILLAGE VOICE, Jan. 17, 1995, *supp.*, at 4. SAT-takers are now graded on a "recentered" curve, which has raised the average score on each section to 500, from 424 for verbal and 478 for math. Recentering is predicted to boost the number of students who score above 740 on the verbal section from 1400 to 10,000.

Even critics acknowledge that ETS has made some improvements in the new test. But preliminary data offers reason to doubt that the new test will substantially reduce concerns about predictive validity, bias, and correlations between parental income and test performance. Although validity data on the new test is not yet available, the College Board has acknowledged that increasing the test's validity was not an objective of the redesign. FAIRTEST, TESTING TITANIC UNDERGOES COSMETIC REDESIGN: "NEW" SAT LAUNCHED ON MARCH 19 (undated published sheet). One study, by University of Georgia psychometrician Stuart Katz, showed that students do not need to read the critical reading passages to score well on the new test. Shteir, *supra*. Predictive pre-test studies as well as preliminary post-test results suggest that the new SAT will not eliminate the test's class bias, racial impact, or underpredictiveness for females. 1995 SAT Scores Rise, ACT Results Flat: Gender, Race Gaps Persist on Both, FAIRTEST EXAMINER, Fall/Winter 1995-96, at 8. "Large differences remain among white males, females and members of most minority groups on both exams." *Id.*

The data presented *infra* at note 152 on the correlation between SAT scores and parental income are based on the results of the new SAT.

68. See Lani Guinier, Michelle Fine, and Jane Balin, *Becoming Gentlemen: Women's Experiences at One Ivy League Law School*, 143 U. PA. L. REV. 1, 23 n.70, 27 n.74 (1994) (showing a weak relationship between LSAT and first-, second-, and third-year grades for men and women, students of color, and white students); Memorandum from Jane Balin, Assistant Professor of

found weak correlations between LSAT scores and first-year law school performance.⁶⁹ For black students, the relatively weak correlations between Index scores and first-year grades were achieved only "if lower undergraduate grade-point averages are made to predict higher first-year averages in the Law School."⁷⁰

Testing experts and decision makers acknowledge that the correlation between test scores and first year college performance for those who actually matriculate is relatively low.⁷¹ However, they point out that the group of people who are accepted by and attend a particular institution come from a narrow subset of the total applicant pool.⁷² To deal with this problem,⁷³ experts have developed a formula that theoretically

Sociology/Anthropology, Colgate University, to Lani Guinier, Paragraph on Our LSAT Findings 1 (July 14, 1995) (on file with authors) [hereinafter Balin Memorandum]:

When analyzing the predictive value of LSAT in terms of law school gpas we found that while significant[,] LSAT is not a very strong indicator of how a student will perform at Penn's law school. . . . In bivariate regressions where LSAT was regressed on law school gpas for years one, two and three, LSAT explains only 21% ($R^2=.21$ $p=.0001$) of the variance in student performance by year three, and even less in years one and two—14% ($R^2=.14$ $p=.0001$) and 15% ($R^2=.15$ $p=.0001$) respectively.

69. Declaration of Martin M. Shapiro at 15, *Hopwood v. Texas*, 861 F. Supp. 551 (W.D. Tex. 1994), (No. A-92-CA-563-SS) (containing table compiled from results of the analyses for the University of Texas School of Law showing the correlation of LSAT with first-year averages (FYA) for white students to be 0.24; when regressed to account for the variance in actual first-year grades, the correlation with LSAT alone equals about 6% (or 0.24-squared)), *rev'd on other grounds*, 78 F.3d 932 (5th Cir.), *cert. denied*, 116 S. Ct. 2582 (1996).

70. *Id.* at 16. In his declaration, Professor Shapiro observes that the use of the regression equation for African American students at the University of Texas School of Law "is invalid." *Id.* at 14.

[T]he multiple correlation between the first-year grade point average predicted by the combination of LSAT and UGPA [undergraduate grade point average] with the actual FYA [at the University of Texas School of Law] is rather poor for African American students, 0.28. A multiple correlation of 0.28 corresponds to the finding that only 8% (0.28-squared) of the variance in the actual FYAs of African American students is accounted for by the multiple regression equation which predicts FYA from LSAT and UGPA.

Id. at 15. What is more significant, however, is that "the multiple correlation coefficient of 0.28 for African American students is achieved only if the UGPA of each African American student is multiplied by a weight equal to -3.35." *Id.* at 15-16. In other words, "for African American students, this small 8% predictability is achieved only if lower undergraduate grade-point averages are made to predict higher first-year averages in the Law School." *Id.* at 16.

71. See Bruce Weber, *Inside the Meritocracy Machine*, N.Y. TIMES, Apr. 28, 1996, §6 (Magazine), at 44, 48 (documenting randomness of the selection process as described by Harvard admissions officers). In the words of one admissions officer, "Harvard could probably admit a whole other class. . . . Admit the class, put it aside and go ahead and admit a whole other one that would do extraordinarily well and go on to do fabulous things in life." *Id.*

72. See CAMERON, *supra* note 42, at 5; Selmi, *supra* note 8, at 1266. For example, a law school may admit applicants with test scores over the 90th percentile. Within this pool, test scores have very little predictive value because they do not correlate well with differentials in first year grades. But testing experts would argue that this does not tell us about the correlation between grades and test scores for applicants with test scores in the 80th percentile.

73. It is impossible to obtain actual data about the correlation between test scores and performance for those who were rejected or did not enroll at the institution. We will never know how they would perform because they never obtained first year grades. The technical term for this problem is "restriction of range." See FAIRNESS IN EMPLOYMENT TESTING, *supra* note 40, at 124-27.

enables a projection from the performance of the known group of applicants to the performance of those who were not admitted and did not enroll.⁷⁴ This mathematical correction increases the correlation between the test and performance in the position.⁷⁵

However, the empirical basis for this mathematical correction has been seriously questioned.⁷⁶ The formula assumes continuity and similarity among applicants so that one can simply project out in a linear fashion from the performance of those who complete their freshman year.⁷⁷ But in many situations, we cannot know, for example, whether the pattern of performance of those who were admitted from the 90th percentile of test-takers would resemble the pattern of those who were not admitted from the 70th or 80th percentile of test-takers.⁷⁸ Some tests predict high test-scorers' performance better than that of lower test-scorers; others offer reliable information only about clear failures.⁷⁹ Yet, the mathematical adjustment assumes that the tests are equally predictive at different levels of performance on the test. Finally, this mathematical formula fails to take into account the many variables other than test scores and grade point average that might influence acceptance, enrollment, and completion of freshman year.⁸⁰ For these reasons, in some contexts experts have taken the position that "the

74. *Id.* at 125 (describing the mathematical formula identifying the 'restriction ratio'). The formula estimates the performance of those outside the pool of admittees, projecting out from the performance of those who did enroll.

75. See Selmi, *supra* note 8, at 1266; Kelman, *supra* note 8, at 1213 n.157:

The purpose of range restriction is illustrated by a simple analogy. If a law school admits only those applicants scoring in a very narrow, high range on the LSAT, we might not be surprised to find that the small differences in LSAT scores among that quite similar group did a poor job predicting differences in grades. It would still be true, however, that if the school admitted some students with much lower LSAT scores, those students would get lower grades. Thus, to get a more accurate measure of [the correlation between test scores and performance] here one would "correct" measured validity[,] which is the correlation that actually shows up in the available data, without any adjustments.

76. See CROUSE & TRUSHEIM, *supra* note 43, at 50 ("[T]hese techniques all require assumptions that are difficult to test."); FAIRNESS IN EMPLOYMENT TESTING, *supra* note 40, at 128 ("Lack of adequate reliable data about the variance of test scores in realistically defined applicant populations is a major problem . . ."); Selmi, *supra* note 8, at 1266 (noting a circular quality underlying the correction). As several recent studies have shown, the standard correction formula fails to take account of the fact that many candidates only apply for positions that they could plausibly obtain. So the range of applicants is not random, and is likely to be considerably more restricted than the general population of test-takers. FAIRNESS IN EMPLOYMENT TESTING, *supra* note 40, at 125-26; Kelman, *supra* note 8, at 1213 n.158 (questioning existence of abstract or context-independent answer to question of how to correct for range restriction).

77. FAIRNESS IN EMPLOYMENT TESTING, *supra* note 40, at 125; Selmi, *supra* note 8, at 1266-67.

78. See FAIRNESS IN EMPLOYMENT TESTING, *supra* note 40, at 125.

79. Selmi, *supra* note 8, at 1267. Selmi points out that, if "the admitted group would have performed better than the observed relationship predicts then including that group in the analysis may decrease rather than increase the observed correlation." *Id.*

80. CROUSE & TRUSHEIM, *supra* note 43, at 49.

conservative response is to apply no correction for restriction of range."⁸¹

Indeed, empirical and statistical evidence suggests that many of those who are excluded based on test results could perform comparably to those admitted. Many tests exclude applicants who could in fact perform successfully. A vivid example arose from an error in the scoring of the 1976 version of the Armed Services Vocational Aptitude Battery. A calibration error resulted in the admission to the military of over 300,000 recruits who actually failed the screening test used by the armed services.⁸² Studies examining the subsequent performance of those "potentially ineligible" ("PIs") found that performance differentials were "not large and in several cases the PIs performed as well as or even better than the controls."⁸³ The PIs "completed training; their attrition rates weren't unusually high; they were promoted at rates only slightly lower than their higher-scoring peers; and they reenlisted."⁸⁴

It is widely recognized that high school grades are more predictive of college freshman-year grades than the SAT.⁸⁵ Perhaps even more significant is the extremely small increase in predictiveness gained by using the SAT in conjunction with high school grades. "Colleges now make only trivial improvements in their ability to forecast college completion correctly when they use the SAT and rank together instead of using rank alone."⁸⁶ The incremental value of the SAT in predicting

81. FAIRNESS IN EMPLOYMENT TESTING, *supra* note 40, at 128 (emphasis omitted); *see also* CROUSE & TRUSHEIM, *supra* note 43, at 49-50 (expressing skepticism about whether corrections used to address range restriction are warranted and make improvements over uncorrected predictions and estimates).

82. WALTER M. HANEY ET AL., THE FRACTURED MARKETPLACE FOR STANDARDIZED TESTING 46 (1993).

83. *Id.*

84. *Id.* at 46-47.

85. *See* WILLINGHAM ET AL., *supra* note 66, at 11 (finding that the school record provides a more accurate forecast of FGPA [freshman grade point average] than do test scores (.48 vs. .42)); *see also* CAMERON, *supra* note 42, at 5 (same); CROUSE & TRUSHEIM, *supra* note 43, at 40 (quoting the Educational Testing Service ("ETS"), the authors of the SAT, as saying that "students' previous grades are the most important indicators . . ."); Christopher Lee, *Schools Question SAT's Use*, DALLAS MORNING NEWS, July 25, 1995, at 1A (quoting College Board spokesperson as stating that grades are a better predictor, but noting that the combination of SAT scores and grades improves predictive validity).

86. James Crouse, *Does the SAT Help Colleges Make Better Selection Decisions?*, 55 HARV. EDUC. REV. 195, 209 (1985) ("[C]orrect forecasts increase only 0.1 per hundred by using the SAT with the 2.5 grade point average (GPA) admissions standard and by 0.2 per hundred using the 3.0 GPA admissions standard. . ."). Furthermore, "the added prediction value of the SAT over high school rank alone is never more than 1.3 fewer errors per 100 admissions decisions . . . for either [freshman grade point average] or bachelor's degree outcomes." CROUSE & TRUSHEIM, *supra* note 43, at 58; *see also* CAMERON, *supra* note 42, at 5 (noting that even for those colleges with standard deviations of 100 or higher, the average correlation of SAT with GPA is .49, as compared with the median correlation of high school record and freshman GPA of .48).

bachelor's degree attainment is even smaller.⁸⁷ Similar questions have been raised about whether cognitive tests used in the employment context have greater predictive value than alternative screening devices "such as education, job experience, peer evaluations, and interviews."⁸⁸

It is also difficult to justify the use of rankings to distinguish at the margins among people whose performance falls within a relatively narrow band. Our yardsticks of merit can be used to differentiate yards, perhaps, but not inches or half-inches. The statistical concept of the standard error of measurement suggests that any particular test score is only indicative of an individual's "scoring range."⁸⁹ Although tests may offer useful information about people at the very top and very bottom of the pool, they cannot reliably differentiate among candidates at the margins or in narrow bands of test scores.⁹⁰ Yet many "reverse discrimination" cases involve candidates whose scores fall within relatively narrow bands that cannot reliably be distinguished.⁹¹

87. CROUSE & TRUSHEIM, *supra* note 43, at 58 ("correct forecasts increase only 0.1 per 100 by using the SAT with the 2.5 predicted GPA admission standard, and by 0.2 per 100 using the 3.0 GPA admissions standard"). The results are much the same regardless of where the admissions cutoff point is set. *Id.* at 55. See also CAMERON, *supra* note 42, at 5 (even for those colleges with standard deviations of 100 or higher, the average correlation of SAT with GPA is .49, as compared with the median correlation of high school record and freshman GPA of .48).

88. Kelman, *supra* note 8, at 1212.

89. Selmi plays out this concept in the context of the facts of *Johnson v. Transportation Agency*, 480 U.S. 616 (1987):

The standard error of measurement . . . provides the necessary analytical tool to determine whether Paul Johnson's test score of 75 was meaningfully different from Diane Joyce's score of 73. Assuming a standard error of 5, and a 95% confidence level, we can be reasonably certain that on subsequent examinations Paul Johnson would score between 65 and 85 and Diane Joyce would score between 63 and 83. We have little confidence, however, in predicting who will perform better on any particular examination.

Selmi, *supra* note 8, at 1274 (footnote omitted). As George Hanford, the President of the College Board of Examiners from 1979 to 1987, put it: "The SAT's error of measurement is such that two times out of three, the score a student gets on a particular form, or edition, of the test will be within about 30 points one way or the other of the true score." GEORGE H. HANFORD, *LIFE WITH THE SAT: ASSESSING OUR YOUNG PEOPLE AND OUR TIMES* 33 (1991).

90. David B. Wilkins and G. Mitu Gulati explain this phenomenon in the context of law firms:

[A]ssuming that these firms face a normal bell-shaped distribution of worker talent (that is, a small number of "outstanding" and "unacceptable" workers at either end with the vast majority of candidates clustered together in the "average" range), they should be relatively indifferent as to which average candidates are hired. Since quality is subjective and therefore difficult to evaluate, the "signals" applicants use to demonstrate their "merit" (for example, educational credentials, recommendations, work experience) will inevitably be "noisier" (that is, less reliable predictors of actual quality) the closer one gets to the mean.

David B. Wilkins & G. Mitu Gulati, *Why Are There So Few Black Lawyers in Corporate Law Firms?*, 84 CALIF. L. REV. 493 (1996).

91. Selmi, *supra* note 8, at 1275 (finding that measurement theory would indicate that two individuals in the same relatively narrow band would be equally qualified, despite a difference in test scores); see also Kelman, *supra* note 8, at 1213 nn.157-58 (finding that small differences in LSAT scores among those actually selected do not provide reliable basis for predicting differences in grades); Selmi, *supra* note 8, at 1266 (same).

An additional problem in establishing correlations between test scores and performance stems from the possibility that some people who may perform well in an educational or work environment perform poorly under the unique circumstances of most testing conditions.⁹² Moreover, most test instruments measure a wide range of skills and abilities through the narrow lens of the linguistic and logical-mathematical domains; if test-takers "are not strong in those two areas, their abilities in other areas may be obscured."⁹³ There is developing evidence that attributes that cannot easily be measured through standardized paper-and-pencil tests, such as discipline, emotional intelligence, commitment, drive to succeed, and reliability, may be more important to successful work or school performance than marginally better performance on tests of general intelligence or analytical ability.⁹⁴ Emotional qualities, what Daniel Goleman calls "emotional intelligence," may be just as important a predictor of academic success as test-taking ability.⁹⁵ Standardized tests "do not measure motivation, perseverance or teamwork skills."⁹⁶

A study of three classes of Harvard alumni over three decades, for example, found a high correlation between "success"—defined by income, community involvement, and professional satisfaction—and two criteria that might not ordinarily be associated with Harvard fresh-

92. See Claude M. Steele & Joshua Aronson, *Stereotype Threat and the Intellectual Test Performance of African Americans*, 69 J. PERSONALITY & SOC. PSYCHOL. 797, 808 (1995) (finding that a differential in performance between black and white candidates with identical SAT scores was attributable to stereotype threat, a concern with the significance of one's performance in light of a devaluing stereotype). These individuals may succumb to the distorted stress of the test-taking environment, which does not reproduce in any meaningful way the stress of the work environment.

93. GARDNER, *supra* note 21, at 10, 176.

94. *Id.* at 166-73 (questioning the relationship of performance on standardized tests to performance at work or in other problem-solving activities, and describing limited predictiveness of creativity tests and shift among researchers to examining what actually happens when individuals engage in problem-solving or problem-finding activities); *id.* at 31 (observing that paper-and-pencil tests sample only a small proportion of intellectual abilities and often reward a certain kind of decontextualized facility); DANIEL GOLEMAN, *EMOTIONAL INTELLIGENCE* 35 (1995) [hereinafter GOLEMAN, *EMOTIONAL INTELLIGENCE*] (citing studies showing that the ability to handle frustrations, control emotions, and get along with others has greater significance in achieving success, and finding that men with the highest test scores in college were not particularly successful compared to their lower-scoring peers in terms of salary, productivity, or status in their field); Daniel Goleman, *The Decline of the Nice-Guy Quotient*, N.Y. TIMES, Sept. 10, 1995, §4, at 6 (discussing a recent study at Bell Laboratories, a high-tech "think tank," which found that the most valued and productive engineers "were not those with the highest IQs, the highest academic credentials or the best scores on achievement tests," but those who "excelled in rapport, empathy, cooperation, persuasion and the ability to build consensus among people").

95. See GOLEMAN, *EMOTIONAL INTELLIGENCE*, *supra* note 94, at 34.

96. Steven A. Holmes, *A Rage for Merit, Whatever That Is*, N.Y. TIMES, July 30, 1995, §4, at 6 (quoting Alexandra Wigdor, director of the Division of Education, Labor and Human Performance of the National Academy of Science).

men: low SAT scores and a blue-collar background.⁹⁷ When asked what predicts life success, college admissions officers at elite universities report that, above a minimum level of competence, "initiative" or "hunger" are the best predictors.⁹⁸

Success may simply reflect a person's opportunities to learn a job or a skill, opportunities that are not measured by any paper-and-pencil test.⁹⁹ One-size-fits-all testing may also compromise the institution's capacity to search for what it really values in selection. Researchers are becoming increasingly aware that privileging the aspects of performance rewarded by standardized tests may well screen out the contributions of people who would bring important and different skills to the workplace or educational institution.¹⁰⁰ Finally, individual performance in both the workplace and educational environments is often enhanced when challenged by competing perspectives or when given the opportunity to develop in conjunction with the different approaches or skills of others.¹⁰¹

The problem of using standardized tests to predict performance is even more acute in the employment context. Standardized tests may reward qualities such as willingness to guess, conformity, and docility.¹⁰² If this is so, then test performance may not relate significantly to the capacity to function well in jobs that require creativity, judgment, and leadership. In a service economy, creativity and interpersonal skills are important, though hard to measure. In the stock scenario of civil service exams for police and fire departments, traits such as honesty, courage, and ability to manage anger are left out.¹⁰³ In other words, people who rely heavily on numbers to make employment decisions "are being misled."¹⁰⁴

97. David K. Shipler, *My Equal Opportunity, Your Free Lunch*, N.Y. TIMES, Mar. 5, 1995, §4, at 1, 16.

98. Marlyn McGrath Lewis, director of admissions for Harvard and Radcliffe, said, "We have particular interest in students from a modest background. Coupled with high achievement and a high ambition level and energy, a background that's modest can really be a help. We know that's the best investment we can make: a kid who's hungry." *Id.*

99. See *infra* notes 221-226 and accompanying text (describing the central role of opportunity to learn skills or perform tasks).

100. See GARDNER, *supra* note 21, at 15-34 (describing multiple intelligences and the limitations of standardized tests in identifying them).

101. We develop this argument both in the context of "functional diversity," *infra* Part IV.C.1, and in terms of new approaches to assessment, *infra* notes 215-217, 236-241, and Part IV.A.

102. Selmi, *supra* note 8, at 1265; see also GARDNER, *supra* note 21, at 168 (noting that individuals with well-developed skills of abstraction can "psych out" standardized tests, scoring well when they know little about what is being tested).

103. Holmes, *supra* note 96, at 6 (quoting James Heckman, professor of economics and public policy at the University of Chicago, as saying that "[t]hose kinds of private skills are totally missed in the merit argument").

104. *Id.* (quoting Benjamin Wright, professor of psychology and education at the University of Chicago).

Thus, if John Doe scored higher on the civil service exam, that does not necessarily mean that he would perform better as a police officer. Cheryl Hopwood's higher Texas Index score does not necessarily establish her superior qualifications to be a law student (or a lawyer). Yet, many employers and schools effectively use test scores to rank candidates or establish cut-off scores within relatively narrow bands of test performance.¹⁰⁵ Colleges and professional schools often use cut-off scores or presumptions to create categories of students who will be automatically admitted or rejected.¹⁰⁶ Small, statistically insignificant differences in test scores may well determine whether a candidate is admitted or rejected.¹⁰⁷ Schools and employers use these tests even while they admit that the tests have very limited predictive value. Similarly, complainants in cases challenging affirmative action policies also rely on their test scores, even though those scores may not, in statistical terms, tell us very much about their likely performance in a job or in law school.¹⁰⁸

Advocates of objective tests usually respond that, although these tests are limited, they are the best we can do. One of the most common and facially compelling arguments for retaining uniform aptitude tests, such as the SAT, focuses on their role as a "leveling agent":

In addition to its supplemental and incremental value, the SAT has a value of its own in confirming the grades from different schools. Since the high school record is a reflection of locally controlled curricula and local grading practices . . . there are variations from school to school in the meaning of the

105. Selmi, *supra* note 8, at 1253 (citing examples of employment decisions based on small point score differentials, and concluding that "any test score difference is assumed to imply that the selection was antimeritocratic").

106. See FAIRTEST, WHAT'S WRONG WITH THE SAT? (undated published sheet) ("[C]olleges and agencies such as the National Merit Scholarship Corporation routinely use cut-off scores, where even 10 points—just one question—can mean the difference between acceptance and rejection."); ETS Developing "New" GRE, FAIRTEST EXAMINER, Fall/Winter 1995-96, at 10, 11, 13 (finding that 27% of graduate schools that require the GRE reported that they use a cutoff, and 50% reported using test scores to determine fellowship awards); see also Hopwood v. Texas, 78 F.3d 932, 935 (5th Cir.) (alleging that the Texas Index's administrative usefulness was its ability to sort candidates into categories of presumptive admit, presumptive deny, and the discretionary zone), *cert. denied*, 116 S. Ct. 2582 (1996). The University of Texas, like many other law schools, weights the LSAT more heavily than grades. This may result in part from the formula used by *U.S. News and World Report* to rank law schools.

107. See FAIRTEST, SAT MISUSE (undated published sheet) ("Despite the obvious unfairness of using the SAT—which has a 68 point margin of error—in such a precise way, over 400 schools . . . require minimum test scores.").

108. Selmi, *supra* note 8, at 1275; see also St. George, *supra* note 2, at A17 (reporting that Chicago city lawyers argued, in response to complaints about a city affirmative action plan, that "test experts say even the best-designed tests don't predict job performance—so that someone scoring 94 does not necessarily make a better lieutenant than someone with a 90").

grades. . . . The SAT, on the other hand, represents a standardized measure of the same mental tasks. . . .¹⁰⁹

This argument is often supplemented by concerns that eliminating the SAT would induce high schools to give more A's.¹¹⁰ Thus, standardized tests are offered as a solution to the problem of grade inflation and nonuniformity.

However, this stated concern about the difficulty of interpreting local differences in grading practices does not account for the practices actually used to assess high school records. Colleges often improve their capacity to evaluate high school performance by relying on rank in class, rather than simple grade point average, and can demand that high schools supply grade distributions for their entire graduating classes.¹¹¹ In addition, colleges and professional schools have relationships and track records with institutions that enable them to make informed judgments.¹¹²

More importantly, the data undercuts the significance of the "leveling" argument as a basis for relying on standardized tests to improve selection decisions. For purposes of selection, grade inflation and nonuniformity in grading practices only matter if they undercut the value of high school grades in predicting performance in college. The evidence suggests exactly the opposite. Whatever the variability in local grading practices, high school grades have consistently proven more predictive of freshman grade point average than the SAT.¹¹³ This has remained true for both selective and non-selective colleges.¹¹⁴ Indeed, studies show that during the period of 1981-88, the predictive value of high school grades has increased for more selective institutions, while the predictive value of SAT scores has decreased slightly.¹¹⁵ During a period in which high schools may have experienced grade inflation, there was no detectable decline in the predictive value of high school

109. CROUSE & TRUSHEIM, *supra* note 43, at 134-35 (quoting Thomas F. Donlon and William H. Angoff, *The Scholastic Aptitude Test*, in THE COLLEGE BOARD ADMISSIONS TESTING PROGRAM: A TECHNICAL REPORT ON RESEARCH AND DEVELOPMENT ACTIVITIES RELATING TO THE SCHOLASTIC APTITUDE TEST AND ACHIEVEMENT TESTS 15-16 (William H. Angoff ed., 1971)).

110. *Id.* at 70-71.

111. *Id.* at 71.

112. *Id.* at 147.

113. WILLINGHAM ET AL., *supra* note 67, at 24-25 (showing from the mid-1960s to 1988 a consistently higher correlation between high school record and freshman GPA than between SAT and freshman GPA).

114. See CROUSE & TRUSHEIM, *supra* note 43, at 143-44 (finding that "the high school record is the best single predictor of college grade point average," and that "controlling selectivity has only a trivial impact" on college GPA); WILLINGHAM ET AL., *supra* note 66, at 56-57 (finding that, starting with the 1977-80 period, high school record has shown an equal or higher correlation with freshman GPA at both selective and non-selective colleges).

115. WILLINGHAM ET AL., *supra* note 66, at 57.

grades.¹¹⁶ During this same period, the SAT did not appreciably improve colleges' capacity to select those most likely to perform well in their freshman year.¹¹⁷

Another important justification for relying on standardized tests stems from their cost-effectiveness. Standardized tests can be administered to huge numbers of applicants at relatively low cost. This view of "cost-effectiveness" focuses on short-term expenses of selection. It fails to take into account the costs to institutions of using selection criteria that do not predict successful performers.¹¹⁸ It denies institutions that use tests to screen applicants the capacity to admit potential leaders from unconventional backgrounds, whose skills are not easily quantifiable.¹¹⁹ It also fails to consider the full range of costs incurred in developing and relying predominantly on standardized tests to predict performance. These costs include the social consequences of excluding those who have the capacity to succeed but are "stigmatized" by their weak performance on the tests.¹²⁰ They also include the costs of failing to explore more innovative and potentially productive ways of selecting candidates. As we show in the next Section, it also tends to give the wealthy a thumb on the scale of merit.

B. Standardized Testing and the Fiction of Fairness

The previous Section challenges the assumption, implicit in the stock affirmative action narratives, that those who score higher on stan-

116. CROUSE & TRUSHEIM, *supra* note 43, at 136.

117. *Id.* at 136-37.

118. See *infra* notes 277-278 and accompanying text.

119. See Hugh Price, President, National Urban League and former Vice President, Rockefeller Foundation, *Toward An Inclusive Society*, Speech to the Commonwealth Club of California (Feb. 10, 1995), in *COMMONWEALTH*, Feb. 27, 1995, at 1, 4:

In the spring of 1963, I really butchered the law boards, which was the admissions test back then. Yet Yale admitted me, even though my score was probably 200 points below that of the average white enrollee. Though an A student in high school, I'd been a solid B student at Amherst, but surely no academic superstar.

Despite that miserable test score, my grades at Yale never scraped absolute bottom. If memory serves, I finished toward the lower end of the middle third of our class. Since we had only seven blacks out of 150 students in my class, that obviously means there were many whites [sic] students with higher law board tests beneath me in the rankings.

At my 20th reunion, several Yale classmates said I was one of the "stars" of our class.

Id. at 4.

120. Short-term cost effectiveness is a narrow and potentially dysfunctional view of efficiency. It externalizes the social costs of unnecessarily limiting access to opportunity. By camouflaging the weakness of prediction in general, it exaggerates the amount of confidence we should place on such measures. Moreover, performance on predictors is valorized as a measure of self-worth. Cf. Lynda Richardson, *Dr. Crew's Prescription: 'Efficacy' Looms as New York's Next Education Philosophy*, N.Y. TIMES, Nov. 26, 1995, at 39, 42 (quoting Dr. Jeffrey Howard, whose "efficacy" program focuses on each student's opportunity for continuous improvement: "Americans are taught that half of our children are below average in intelligence. It's a crazy and limited idea that results in the destruction of self-confidence and their elimination from real learning opportunities.").

dardized tests will function better in the positions they seek. This Section takes on the second implicit assumption of those narratives: that the conventional methods of selecting candidates for high-stakes positions are fair. To assess the implicit claim that existing selection criteria are fair, it is crucial to broaden the conception of fairness that frames the analysis.¹²¹ The stock affirmative action narrative implicitly embraces process definitions of fairness: do applicants receive the same treatment in the evaluation process? Are employers adhering to the stated standards for everyone? This concept of fairness is misleading. It presumes a level playing field—that if everyone plays by the same rules, the game does not favor or disadvantage anyone.

There is, however, a conception of fairness, which we call “fairness as equal access and opportunity,” that summons the substantive dimension beneath formal sameness. This more substantive conception challenges the assumption that in all situations sameness equals fairness. It focuses on providing members of various races and genders equivalent opportunities to demonstrate their capacities. It focuses attention on the fact that formal sameness camouflages actual difference.¹²² And it focuses attention on differential access and the exclusionary bias built into the screening devices used to allocate positions.

This conception of fairness requires that the standards governing the process not arbitrarily advantage members of one group over another. If different approaches can be used to accomplish an underlying goal, and these differing approaches correlate with race or gender, then an employer may not fairly insist on only one of these approaches.¹²³

121. See Fish, *supra* note 32, at 136:

The sleight-of-hand logic that first abstracts events from history and then assesses them from behind a veil of willed ignorance gains some of its plausibility from . . . the invocation of fairness[, which] is used to legitimize an institutionalized inequality. . . . Bizarre as it may seem, individualism in this argument turns out to mean that everyone is or should be the same. This dismissal of individual difference in the name of the individual would be funny were its consequences not so serious: it is the mechanism by which imbalances and inequities suffered by millions of people through no fault of their own can be sanitized and even celebrated as the natural workings of unfettered democracy.

Id.

122. See Susan P. Sturm, *Sameness and Subordination: The Dangers of a Universal Solution*, 143 U. PA. L. REV. 201, 210-12 (1994):

[I]nsisting on sameness in the face of difference can itself perpetuate subordination. . . .

....

... Claims of universal sameness allow us to impose norms embedded in the status quo without acknowledging that we are doing so, and without responding to those putting forward a different normative agenda. . . .

....

The commonality approach hides the allocation of power embedded in the ideal of uniformity.

123. Cf. Barbara J. Flagg, *Fashioning a Title VII Remedy for Transparently White Subjective Decisionmaking*, 104 YALE L.J. 2009, 2032-33 (1995) (arguing for a pluralist, equal-access conception of equality challenging employment practices that create unnecessary barriers to employment for minorities); Joel Friedman, *Redefining Equality, Discrimination, and Affirmative*

Alternatively stated, fairness must rest on genuine equality of opportunity.¹²⁴ Fairness means that "[e]veryone should be given an equal opportunity to achieve all that their abilities allow."¹²⁵ It is not "fair" in this sense to use entry-level credentials that appear to treat everyone the same, but in effect deny women and people of color a genuine opportunity to demonstrate their capacities.¹²⁶

In fact, the "testocracy" does not provide a fair playing field for candidates. First, many standardized tests are substantively unfair because they assume that there is a single, uniform way to complete the job, and then tests applicants solely upon criteria consistent with this uniform style. In this way, the testing process entrenches the status-quo mode of production, excluding those individuals who may perform the job just as effectively through different approaches. Second, conventional selection methods advantage candidates from higher socioeconomic backgrounds and disproportionately screen out women and people of color, as well as those in lower-income brackets. When combined with other unstructured screening practices, such as personal connections and alumni preferences, standardized testing creates an arbitrary barrier for many otherwise-qualified candidates.

Action Under Title VII: The Access Principle, 65 TEX. L. REV. 41 (1986) (arguing for an access principle requiring equal opportunity to compete for jobs and promotions based on ability and effort); cf. MARTHA MINOW, MAKING ALL THE DIFFERENCE: INCLUSION, EXCLUSION, AND AMERICAN LAW 16, 84 (1990) ("Strategies for remaking difference include challenging and transforming the unstated norm used for comparisons.").

124. Harris & Narayan, *supra* note 7, at 13. Some commentators have articulated another substantive conception of fairness that emphasizes the opportunity to participate in the process of defining standards for selection that reflect truly representative, collective decision making. Fairness is used as a substantive expression of the legitimacy of the process used to define the existing selection process. Fairness as legitimacy reflects the importance of genuine participation and real consent. In this sense fairness means democratic decision making or the idea that people who feel they have a decision making voice are more likely to accept the ultimate decision, even if it is different from the one they initially supported. This concept of fairness overlaps with a theory of justice, which we address below as part of our articulation of a political theory of citizenship that justifies requiring employers and educational institutions to adopt the least exclusive screening process available to serve express goals. See *infra* Part IV.E.

125. Bass, *supra* note 35; see also Harris & Narayan, *supra* note 7, at 4 (viewing affirmative action as "an attempt to offer . . . [beneficiaries] greater equality of opportunity in a social context marked by pervasive inequalities, one in which many institutional practices work to impede a fair assessment of the capabilities of those who are working class, women, or people of color").

126. It is important to distinguish our proposed conceptions of fairness from those that define fairness as proportionality. Under the proportionality view, fairness means numerical symmetry, which is measured by the correlation between the numbers of people hired or admitted and the demographics of the relevant "pool." See, e.g., Bruce Cain, *Voting Rights and Democratic Theory: Toward a Color-blind Society?*, in *CONTROVERSIES IN MINORITY VOTING: THE VOTING RIGHTS ACT IN PERSPECTIVE* 261, 263 (Bernard Grofman & Chandler Davidson eds., 1992). Cain notes that Americans "prefer to use terms such as *fairness* . . . without explicitly defining them, which causes significant confusion because fairness could in fact mean something other than proportionality." *Id.* at 263.

1. *The Underinclusiveness of One-Size-Fits-All Tests: Sameness Is Not Fairness*

Standardized tests adopt a one-size-fits-all approach to measuring successful performance. In addition to ignoring many of the abilities and skills that are crucial to successful performance,¹²⁷ this insistence on narrow, uniform criteria of success fails to take account of the variety of ways in which successful performance on the job can be achieved. There may be a range of styles and approaches to doing a job, each of which may be effective in some circumstances. Indeed, diversity introduces a variety of job approaches that can complement one another and offer new and potentially more effective styles and strategies.

For example, in many police departments, strength, aggressiveness, and speed are the predominant criteria of selection for police officers.¹²⁸ These characteristics relate to a particular mode and concept of policing focusing on "command presence" and control through authority and force.¹²⁹ If the issue is quick reaction time and physical prowess, some of these qualities, such as speed, may be important. But not everything a police officer does requires quick reaction time. Indeed, in some situations, responding quickly gets police officers and whole departments in trouble.¹³⁰

127. See *supra* notes 92-96 and accompanying text (describing the failure of standardized tests to measure adequately traits such as determination, creativity, reliability, and commitment).

128. See, e.g., THE WOMEN'S ADVISORY COUNCIL TO THE LOS ANGELES POLICE COMMISSION, A BLUEPRINT FOR IMPLEMENTING GENDER EQUITY IN THE LOS ANGELES POLICE DEPARTMENT 30 (1993) [hereinafter BLUEPRINT FOR GENDER EQUITY]. "Recruitment materials reflect a prevailing notion of policing as a rough-and-tumble career for which only 'John Waynes' need apply. The Department gives preference to recruits with backgrounds in mostly male-dominated fields such as security guard and athlete rather than mostly female-dominated fields, such as school teacher and social worker." *Id.* at 24. However, "[s]tudies have shown that 80 to 95 percent of policing involves non-violent or service-oriented activities, and that physical strength is not related to job performance. Yet, physical strength continues to be a central focus of recruitment and training at the LAPD." *Id.* at 30. See also *Thorne v. City of El Segundo*, 726 F.2d 459, 464 (9th Cir. 1983) (finding that plaintiff, to qualify as police officer, had to "display sufficient aggressiveness, self-assuredness or probable physical ability to presently handle herself in stress situations"), *cert. denied*, 469 U.S. 979 (1984); Ruth Colker, *Rank-Order Physical Abilities Selection Devices for Traditionally Male Occupations as Gender-Based Employment Discrimination*, 19 U.C. DAVIS L. REV. 761, 796 (1986) ("[J]urisdictions have chosen test items that emphasize traditionally male-valued speed and strength, ignoring actual job requirements.").

129. See Mary Anne C. Case, *Disaggregating Gender from Sex and Sexual Orientation: The Effeminate Man in the Law and Feminist Jurisprudence*, 105 YALE L.J. 1, 88-89 (quoting studies finding that police tactics overemphasize the paramilitary and the physical, and rely on control through authority).

130. See JEROME H. SKOLNICK & JAMES J. FYFE, ABOVE THE LAW: POLICE AND THE EXCESSIVE USE OF FORCE 95-98, 108 (1993); INDEPENDENT COMMISSION ON THE LOS ANGELES POLICE DEPARTMENT, REPORT OF THE INDEPENDENT COMMISSION ON THE LOS ANGELES POLICE DEPARTMENT 88 (1991) [hereinafter CHRISTOPHER COMMISSION REPORT] ("A corollary of [traditional] culture is an emphasis on the use of force to control a situation, and a disdain for a more patient, less aggressive approach.").

This speed-and-strength standard normalizes a particular type of officer: tough, brawny, and macho.¹³¹ However, studies have begun to show that other modes of policing, such as dispute resolution, persuasion, counseling, and promoting community involvement, are also critical and sometimes superior approaches to policing.¹³² One study of the Los Angeles police department, conducted in the wake of the Rodney King trials, recommended increasing the representation of women on the police force as a critical component of a strategy to reduce the level of police brutality and improve community relations. The study found that women often display a more interactive and engaged approach to policing.¹³³

Similarly, an informal survey of police work in some New York City Housing Authority projects found that, because many women housing authority officers could not rely on their brawn to intimidate potential offenders, they developed a mentoring style with young adolescent males.¹³⁴ The women, many of whom came from the community

131. See CHRISTOPHER COMMISSION REPORT, *supra* note 130, at 88 ("Traditional views concerning the nature of police work in general—that is, that police work is a male-oriented profession with a major emphasis on physical strength—foster a climate in which female officers are discouraged.").

132. *Id.* (criticizing current approaches to policing for underemphasizing interpersonal skills, sensitivity, politeness, and the ability to communicate); BLUEPRINT FOR GENDER EQUITY, *supra* note 128, at 10-11 (noting the consensus among police management, city leaders, and the community for a police department that operates under the principles of community policing, which rest on the foundation of a partnership between police and the community "to resolve problems related to crime, fear of crime, enforcement of laws, and quality of life"); PATRICIA W. LUNNENBORG, WOMEN POLICE OFFICERS: CURRENT CAREER PROFILE 110-11 (1989) (reporting data suggesting that women police officers rely less on violence and more on verbal skills in handling conflict, are less likely to be involved in rule violations, and are more effective in handling female victims of violence); Joseph Balkin, *Why Policemen Don't Like Policewomen*, 16 J. POLICE SCI. & ADMIN. 29, 34 (1988); Sean Grennan, *Findings on the Role of Officer Gender in Violent Encounters with Citizens*, 15 J. POLICE SCI. & ADMIN. 1, 84 (1987) (noting that because women lack a "macho" image, they are "more likely to calm a potentially violent situation and avoid injury to all participants").

133. CHRISTOPHER COMMISSION REPORT, *supra* note 130, at 83-84 ("[F]emale LAPD officers are involved in excessive use of force at rates substantially below those of male officers. . . . The statistics indicate that female officers are not reluctant to use force, but they are not nearly as likely to be involved in use of excessive force," due to female officers' perceived ability to be "more communicative, more skillful at deescalating potentially violent situations and less confrontational."); see also LUNNENBORG, *supra* note 132, at 110-11 (summarizing research on policing styles showing women officers are less aggressive and confrontational; are more gentle and compassionate; depend more on verbal skills; and better handle service and domestic calls, public contacts, and community relations); BLUEPRINT FOR GENDER EQUITY, *supra* note 128, at 7; Case, *supra* note 129, at 87; Lewis J. Sherman, *A Psychological View of Women in Policing*, 1 J. POLICE SCI. & ADMIN. 383, 384 (1973) (predicting that increased use of women police officers would reduce incidents of police brutality and improve police-community relations).

134. Telephone Interview with J. Phillip Thompson, Director of Management and Operations, New York City Housing Authority, 1992-93 (Jan. 25, 1996). Thompson, who is now a professor at Barnard College, recounted that an internal evaluation conducted by the Housing Authority revealed that women housing authority officers were policing in a different, but successful, way. As a result of this evaluation, the authority sought to recruit new cops based on their ability to relate to young

they were patrolling, kept the projects safer because they did not approach the young men in a confrontational way. Their authority was respected because they offered respect.¹³⁵

Yet, women are frequently evaluated by their peers and for promotion based on narrow and underinclusive standards and models of effective work.¹³⁶ The existing culture normalizes only one approach to performance and, in the process, reinforces the capacity of some people to be fairly evaluated and to perform.¹³⁷ A recent study of University of Pennsylvania law students observed a similar phenomenon operating for many women in law school.¹³⁸ Even though men and women may be afforded the same treatment, the study found that women do not participate in class as much as men, and that they are significantly less comfortable speaking with professors outside of class.¹³⁹ The law school may be treating all students the same, but this does not mean that this approach will enable all students to participate, learn, and feel included.¹⁴⁰ Sameness may not be fairness in this context.

people, their knowledge of the community, their willingness to live in the housing projects, and their interest in police work. They also offered free housing to any successful recruit willing to live in the projects.

135. *Id.* (recounting that the women officers showed the young men respect, which was critical to the social status needs of these males; and that the men in turn checked their own behavior, out of mutual respect for the women officers).

136. *Id.*; Case, *supra* note 129, at 85, 87-88 (noting Christopher Commission finding that LAPD training officers criticized female officers for a perceived lack of "stereotypically masculine qualities," such as physical stature and upper body strength, and that female officers "'had a real tough time' achieving acceptance within the Department"); see Colker, *supra* note 128, at 793-94 (noting police and fire departments' resistance to valuing flexibility and balance, two traditionally female traits).

137. For example, the police union objected to the New York Housing Authority's targeted recruitment of women, even though the female officers' approach to policing was successful in reducing crime in the housing complex. The police union insisted on the use of a test that disproportionately screened out the women with the very skills the Authority was seeking. Despite great interest among women living in the projects, the Authority was not allowed to give special test preparation to the women they wanted to recruit. Telephone Interview with J. Phillip Thompson, *supra* note 134. Cf. Jacques Steinberg, *Suit Charges Bias in Tests for Custodians*, N.Y. TIMES, Jan. 31, 1996, at B1 (reporting that the Department of Justice has challenged use of the test given to prospective school custodians in New York City, 92% of whom are white, charging that "the tests do not adequately measure the skills needed to keep a school building in good repair").

138. Guinier et al., *supra* note 68. Employers often rely on measures of presumptive merit, such as law school grades, to allocate job opportunities. But again, this is somewhat tautological since those who get the job are essentially getting the opportunity to learn the job. No one claims you learn how to be a lawyer in law school. But do those who get the opportunity based on law school grades turn out to be better lawyers?

139. *Id.* at 59-80.

140. *Id.* at 72, 73-76 (finding that women require friendliness cues before approaching faculty; that some women law students are less successful negotiating barriers to informal faculty/student interactions; and that while a mentoring relationship positively correlates with institutional success, few female students are mentored by the faculty); cf. Robert E. Fullilove & Philip Uri Treisman, *Mathematics Achievement Among African American Undergraduates at the University of California, Berkeley: An Evaluation of the Mathematics Workshop Program*, 59 J. NEGRO EDUC. 463, 463-78

Indeed, hostility or marginalization within a work or educational environment may account for certain anomalies in reported correlations between test scores and performance for women and people of color.¹⁴¹ For example, women's differential experience both in and out of the classroom, reported at the University of Pennsylvania Law School, may explain why many women who come to the law school with LSAT scores virtually identical to those of the men do not perform as well.¹⁴² Although men and women who enter the Law School possess virtually identical entry-level criteria, by the end of the first year, the men are three times as likely to be in the top 10% and 1.5 times as likely to be in the top 50%. Indeed, when controlled for LSAT (meaning if you take two people with identical LSAT scores), race and gender are better predictors of performance in law school. A white male with an LSAT score identical to that of a white female or black male will do better in law school.¹⁴³ Environmental factors may also explain why there is a higher correlation between SAT scores and performance for black students enrolled in predominantly black colleges than for black students enrolled in predominantly white colleges.¹⁴⁴

The retention and success of new entrants to institutions often depend on expanding or altering the measures of successful performance.¹⁴⁵ But because those institutionalized or structured preferences

(1990) (finding that some black students improve their performance in mathematics by studying in groups).

141. See Kelman, *supra* note 8, at 1171 n.36:

The more serious problem arises when a test is culturally biased in favor of the majority group, but the cultural bias is hidden by the fact that those who are most prone to perform poorly on the test (for example, minority workers) perform poorly on the job as a result of discriminatory in-plant practices . . .

Id. See also Karen De Witt, *Blacks Prone to Dismissal by the U.S.*, N.Y. TIMES, Apr. 20, 1995, at A19 (attributing disproportionately high dismissal rate of black federal employees to bias or lack of cultural awareness, poorly trained supervisors or managers, and a general inability on the part of minorities to work the "old boy network").

142. Guinier et al., *supra* note 68, at 21-32.

143. *Id.* at 23 n.70, 27 n.74; see also Balin Memorandum, *supra* note 68, at 1 ("[W]e found that while the impact of LSAT remains the same when adding College GPA and College rank to the above bivariate regressions, it decreased when we added race and gender to the equation.").

144. CAMERON, *supra* note 42, at 13 (finding that for black students in predominantly white colleges, the correlation between the SAT and freshman GPA was .30; but for black students enrolled in predominantly black colleges, the correlation was .38).

145. See BLUEPRINT FOR GENDER EQUITY, *supra* note 128, at 38 (recommending revision of "procedures for job analysis, selection criteria and written examinations to reflect community policing values," in order to address exclusion and devaluation of women in promotion); R. Roosevelt Thomas, Jr., *From Affirmative Action to Affirming Diversity*, in DIFFERENCES THAT WORK: ORGANIZATIONAL EXCELLENCE THROUGH DIVERSITY 27, 28 (Mary C. Gentile ed., 1994) (arguing that companies' difficulties in diversifying begin not at the recruitment stage but later on, when many minorities and women "plateau and lose their drive and quit or get fired"); *cf.* Case, *supra* note 129, at 73 (advocating the need to reevaluate jobs and the characteristics necessary to succeed in them); MINOW, *supra* note 123 at 15-16 (advocating an openness to institutional redesign as a strategy of inclusion).

camouflage their bias, one-size-fits-all testocracies invite some beneficiaries to believe they have earned their status solely on the basis of objective indicators. These so-called meritocracies also invite beneficiaries of affirmative action to believe exactly the opposite—that they did not earn their opportunity.¹⁴⁶ Affirmative action in this sense perpetuates an asymmetrical approach to evaluation. It allows partial and underinclusive selection standards to proceed without criticism. But those “exceptions,” who bring alternative approaches that do not conform to the traditional ones developed without their participation, are visible evidence of the limitations of one-size-fits-all standards.

Thus, the insistence that sameness is fairness marginalizes the legitimate capabilities and approaches of those who do not conform to the “normal” or traditional attributes of a particular position. Not only does this mono-dimensional approach fail to predict accurately the potential success of applicants, it also unfairly disadvantages some women, people of color, and members of other traditionally marginalized groups. In doing so, it deprives institutions of access to information and insights that could enrich everyone’s capacity to perform effectively.¹⁴⁷

2. *The Reality of the Wealth Preference*

In addition to favoring certain underinclusive standards of performance, the testocracy is skewed in favor of wealthy contestants. Data

146. Because these candidates were visibly selected on the basis of criteria that were not explicitly linked to test score performance, they may suffer stigma. See Jennifer L. Hochschild, *Affirmative Action and the Rumor of Black Inferiority*, BLACK ISSUES HIGHER EDUC., Summer 1995, at 64. Sixty percent of black students at Princeton University claim to have been questioned on their merit because of their race or ethnicity. *Id.* (citing Jessica Hall-Valdez et al., Racial Harassment at Princeton, Princeton University, (1992) (unpublished report); see also Isabel Wilkerson, *Discordant Notes in Detroit: Music and Affirmative Action*, N.Y. TIMES, Mar. 5, 1989, §1, at 1, 30 (reporting that black bass player hired in response to pressure for more racial diversity “would have rather auditioned like everybody else [because] [s]omehow this devalues the audition and worth of every other player”). Yet on balance, surveys refute the claim that affirmative action makes black recipients uncertain about whether they are worthy of a job. Hochschild, *supra*, at 65. The typical survey finds that only one tenth of black faculty think affirmative action “perpetuates the myth of minority and female inferiority.” *Id.* Half of well-educated blacks believe that the success of their organization depends a lot on their work, that they work harder than their peers, and that they do their job much better than their peers. *Id.* Overall, 55% of well-off blacks think affirmative action programs help recipients, and only 4% think such programs hurt recipients. *Id.*

147. The field of medicine offers a compelling example of the consequences of developing standards and measures derived from samples including only white men.

Until very recently, medical theory about heart attacks was based solely on research about men. The theory said that when there was chest pain, nausea, or pain radiating down the left arm, that signaled a heart attack. Abdominal pain was not considered a symptom of heart attack. But, recently, physicians realized that in some women, abdominal pain was a symptom of heart attack and that by not recognizing this symptom, they were missing heart attacks in women. For some women, the symptoms of heart attack simply presented differently than in some men. Recognizing this makes the theory better.

Myra H. Strober, *Feminist Economics: What’s It All About?*, Downing Oration Presented at Melbourne University 7 (Sept. 12, 1995) (transcript on file with *California Law Review*).

indicates that many tests correlate quite closely with parental income. Although 1996 data is not available, earlier studies found a score-income correlation equivalent to or higher than the correlation that the Educational Testing Service (ETS) found between SAT scores and first-year grades.¹⁴⁸ At over 25% of the colleges participating in a 1984 validity study conducted by the ETS, the correlation between SAT scores and family income was larger than the correlation between SAT scores and freshman grades.¹⁴⁹

As we stated earlier, in predicting first-year college grades the SAT offers only a trivial improvement over high school grades alone. The correlation between family income and SAT is nearly four times larger than the incremental improvement in prediction offered by the SAT used in conjunction with high school grades.¹⁵⁰ In other words, the SAT, at the margins, is a better predictor of family income than of first-year college grades.

The linkage between test performance and parental income is consistent and striking. "Average family income rises with each 100-point increase in SAT scores, except for the highest SAT category where the number of cases is small."¹⁵¹ As the following table demonstrates, this correlation between income level and test performance persists within every racial and ethnic group and across gender.¹⁵²

148. CROUSE & TRUSHEIM, *supra* note 43, at 126 (describing the correlation between income and SAT as .3, a figure larger than the correlation between SAT and freshman performance); ALLAN NAIRN ET AL., *THE REIGN OF ETS* 203 (1980) (finding that for the years 1964-65, the correlation between income and test score was .4, while the correlation between freshman performance and test scores ranged between .29 and .37).

149. CROUSE & TRUSHEIM, *supra* note 43, at 125.

150. *Id.*

151. *Id.* at 126.

152. This table was compiled by the authors from raw data provided by the College Entrance Examination Board. See also CAMERON, *supra* note 43, at 11 ("There is a positive correlation between income level and standardized test scores.").

1995 PROFILES OF SAT SCORES BY INCOME AND RACE

		NATIONAL			
INCOME			SAT I-V		SAT I-M
Less than \$10,000			354		415
\$10,000-\$20,000			380		433
\$20,000-\$30,000			405		454
\$30,000-\$40,000			420		468
\$40,000-\$50,000			431		482
\$50,000-\$60,000			440		493
\$60,000-\$70,000			448		502
\$70,000 or more			471		533

		LATIN AMERICAN		PUERTO RICAN		MEXICAN	
INCOME		SAT I-V	SAT I-M	SAT I-V	SAT I-M	SAT I-V	SAT I-M
Less than \$10,000		332	380	320	358	330	386
\$10,000-\$20,000		356	402	346	383	349	403
\$20,000-\$30,000		379	424	371	406	369	420
\$30,000-\$40,000		397	442	382	420	384	431
\$40,000-\$50,000		415	460	395	434	399	446
\$50,000-\$60,000		421	473	408	449	409	456
\$60,000-\$70,000		431	482	412	454	415	458
\$70,000 or more		454	514	424	475	430	478

		WHITE		ASIAN		BLACK	
INCOME		SAT I-V	SAT I-M	SAT I-V	SAT I-M	SAT I-V	SAT I-M
Less than \$10,000		409	460	343	482	320	355
\$10,000-\$20,000		418	459	363	500	337	369
\$20,000-\$30,000		428	471	397	518	352	382
\$30,000-\$40,000		433	478	415	528	362	393
\$40,000-\$50,000		439	488	432	537	375	405
\$50,000-\$60,000		446	498	444	549	382	414
\$60,000-\$70,000		453	506	453	558	385	415
\$70,000 or more		475	533	476	595	407	442

Within each racial and ethnic group, SAT scores increase with income.¹⁵³

Reliance on high school rank alone excludes fewer people from lower socio-economic backgrounds. When the SAT is used in conjunction with high school rank to select college applicants, the number of applicants admitted from lower-income families decreases.¹⁵⁴ This is because the SAT is more strongly correlated with every measure of socio-economic background than is high school rank.¹⁵⁵

153. See Shieler, *supra* note 97, at E16.

154. CROUSE & TRUSHEIM, *supra* note 43, at 128 (finding that this decrease occurs because more applicants "fall below the selection cut-off value."). At the most stringent admissions cutoff, the SAT reduces lower-income acceptance by almost five percent. *Id.*

155. *Id.* at 126; see also Howard Goodman, *Penn Takes Its Pitch to Inner Cities Bearing Promises for Minority Students; A Recruiter Starts a Yearly Trek*, PHILA. INQUIRER, Oct. 20, 1995, at A1 ("Across America, well-to-do students fare better on the Scholastic Assessment Test than those

Other factors that correlate with parental income also figure prominently in the current selection system.¹⁵⁶ The student bodies of the elite undergraduate institutions, which currently constitute the primary feeder schools for post-graduate and professional institutions, consist of a disproportionately high socio-economic group.¹⁵⁷ Students who attend state schools or community colleges frequently are downgraded in the score used to compute their ranking in the admissions process of post-graduate institutions. Indeed, Cheryl Hopwood, the plaintiff challenging the University of Texas Law School's affirmative action plan, attended a community college and a state school, and was downgraded from the presumptive admit to the discretionary category as a result.¹⁵⁸

This devaluation of the grades of students from public institutions may result from several premises. Decision makers may assume that, because these schools accept a wider range of students, they are less rigorous, albeit more affordable. This assumption is reinforced by data showing that many students who work full-time to support themselves take longer to graduate.¹⁵⁹ Decision makers may also assume that the students choose state or community schools because they may be less rigorous, rather than more affordable.¹⁶⁰ Yet, recent research on the City University of New York's twenty-five-year-old open-admission policy suggests that a majority of its students, who are typically poorer than other students and are often black or Latino, eventually graduate and that a higher percentage obtain graduate degrees than is true for some who attend the more academically prominent private universities.¹⁶¹

from low-income homes. In the most recent test, students from families earning \$10,000 to \$20,000 scored an average of 813 (out of 1,600). Those whose families made \$50,000 to \$60,000 scored an average of 933.").

156. See CAMERON, *supra* note 42, at 11.

157. See Lily Eng & Karen Heller, *State Schools Grow Less Affordable*, PHILA. INQUIRER, Apr. 3, 1996, at A1 (noting that private institutions are too expensive for much of the nation, and tuition is becoming unaffordable for middle class families).

158. Hopwood v. Texas, 861 F. Supp. 551, 557 (W.D. Tex. 1994), *rev'd on other grounds*, 78 F.3d 932 (5th Cir.), *cert. denied*, 116 S. Ct. 2582 (1996).

159. See Karen W. Arenson, *Study Details Success Stories in Open Admissions at CUNY*, N.Y. TIMES, May 7, 1996, at A1, B4 (citing research from DAVID E. LAVIN AND DAVID HYLLEGARD, CHANGING THE ODDS: OPEN ADMISSIONS AND THE LIFE CHANCES OF THE DISADVANTAGED (1996), that students who attend community college or city university take longer to graduate—sometimes more than a decade longer).

160. But see *Survival Guide to College Costs*, PHILA. INQUIRER, Mar. 31, 1996, at A27 ("State colleges and universities offer little aid but the cost of attending is generally half that of a private institution. Many students are applying to both, . . . then seeing what the ultimate cost will be once aid packages are determined.").

161. *Id.* Indeed, a National Research Council study of doctorates showed that from 1983 to 1992, 3877 City University graduates—about 15% of the total for all of New York State—went on to earn Ph.D.'s. Brooklyn College, a City University, provided 947 eventual doctorates while New York University, a prestigious private school nearly 50% larger than Brooklyn College, produced only 934. *Id.*; see also *City College Praises Immigrant Achievers*, N.Y. TIMES, June 5, 1996, at B10 (discussing two valedictorians at New York's Hunter College who achieved perfect 4.0 grade point

Thus, Cheryl Hopwood may well be the victim of a class bias in the admissions process that mirrors the bias confronting applicants of color.

Children of higher-income parents have another clear advantage over their poorer competitors. They can afford to pay for coaching in both test-taking and the packaging of their accomplishments. This extra help includes engaging consultants to advise and assist students in presenting a strong application.¹⁶² It also includes test preparation courses, many of which cost between \$500 and \$1000, that coach future test-takers in the art of test-taking.¹⁶³ Studies suggest that coaching can be effective in increasing test scores.¹⁶⁴ Of particular interest are studies showing that race and parental income "were the most significant predictor variables contributing to the verbal individual coaching effect."¹⁶⁵ One study found that the average verbal coaching effect for black students was 46.7 points above that for white students.¹⁶⁶

What is often touted as a merit-based standard is instead arbitrary and exclusionary. It is arbitrary when it does not correlate well with what it is supposed to be measuring.¹⁶⁷ It is exclusionary when test scores often correlate so reliably with parental income.¹⁶⁸ Test scores tell us more about the past than the future. Thus, the move to objective tests minimized the visibility of one type of bias and class privilege, but the bias reemerged in less obvious, but in some ways more pervasive, forms. Sameness is not necessarily fairness. Yet, the stock affirmative action

averages: one took eight years to complete his diploma, while the other began at Hunter taking remedial classes).

162. See Emily M. Bernstein, *College Matchmakers: Consultants Plot Ways to Impress the 'Right' School*, N.Y. TIMES, Apr. 17, 1996, at B1.

163. See Goodman, *supra* note 155 (finding that a poorer student is less likely to attend well-funded schools, to be able to afford SAT preparation courses, to take the test more than once, and to come from a household where education is cherished).

164. One study found that training for the SAT caused an average gain of 29 points for the SAT-Verbal and 33 points for the SAT-Math. Warner V. Slack & Douglas Porter, *The Scholastic Aptitude Test: A Critical Appraisal*, 50 HARV. EDUC. REV. 154, 160 (1980). Another study found average increases of 110 points over "normal growth" as the result of test coaching. SARAH STOCKWELL ET AL., THE SAT COACHING COVER-UP 11 (1991).

165. Beverly P. Cole, *College Admissions and Coaching*, NEGRO EDUC. REV., Apr.-Jul. 1987, at 125, reprinted in TESTING AFRICAN AMERICAN STUDENTS 97, 103 (Asa G. Hilliard III ed., 2d ed. 1995) (reporting results from the Stroud "Re-Analysis of the FTC Study of Commercial Coaching for the SAT").

166. *Id.* Another study of the General Aptitude Test Battery (GATB) found that the test would be vulnerable to coaching. FAIRNESS IN EMPLOYMENT TESTING, *supra* note 40, at 116.

167. The managing partner at a large New York law firm confirmed this when he described a study his firm conducted to assess which hires eventually became partners over a thirty-year period. Conversation between Robert Preiskel, former managing partner at Fried, Frank, Shriver & Jacobsen, and Lani Guinier (Mar. 1993). The study found that those who were superstars in law school were also likely to be outstanding lawyers and to become partners. Those who were the top 1% in law school were often the top 1% of the firm's lawyers. But below that top 1%, the results were often random: there was little correlation between law school grades and actual performance in the profession.

168. See *supra* notes 152-153 and accompanying charts.

narrative hides these connections among test-centered merit, fairness-as-sameness, and socio-economic status. Cheryl Hopwood and John Doe do not see the class-based connection between their own exclusion and that of the beneficiaries of affirmative action.

3. *The Reality of Racial and Gender Bias*

There is no dispute that existing methods of selection, both objective and subjective, exclude people based on their race and gender. Gender and race differentials in standardized tests have been widely documented, both in the scholarly literature and in litigation.¹⁶⁹ Thus, reliance on these devices for determining merit screens out a disproportionate number of women and people of color who apply for positions.

169. The following chart from 1994 illustrates the gaps in performance on the SAT by race, ethnicity, and gender:

	Verbal		Mathematics	
	Men	Women	Men	Women
American Indian	397	395	459	425
Asian American	417	414	557	514
Black	348	354	399	381
Mexican American	375	368	448	410
Puerto Rican	371	365	432	395
Other Hispanic	389	379	462	414
White	445	441	519	475
Other	426	424	505	459
National averages	425	421	501	460

Gender Gap Continues to Close on S.A.T.'s, N.Y. TIMES, Aug. 25, 1994, at A12 (quoting data taken from the College Board); see also PHYLLIS ROSSER, CENTER FOR WOMEN'S POLICY STUDIES, THE SAT GENDER GAP: IDENTIFYING THE CAUSES 4 (1989) ("In 1988, women's average SAT scores were 56 points lower than men's: 13 points on the Verbal Section . . . and 43 points on the Math Section. However, the College Board's own Validity Studies show that women's average first year college grades are as good or better. . . ."); Robert B. Slater, *Ranking the States by Black-White SAT Scoring Gaps*, J. BLACKS HIGHER EDUC., Winter 1995/1996, at 71 (documenting that blacks on average score 110 points below whites on the math portion of the SAT and 92 points below whites on the verbal portion).

Other sources documenting similar phenomena include: FAIRNESS IN EMPLOYMENT TESTING, *supra* note 40, at 20 (reporting "significant group differences in average test scores, which have been demonstrated with virtually all standardized tests. Blacks as a group score well below the majority group and Hispanics fall roughly in between as a rule."); OFFICE FOR MINORITY EDUCATION, EDUCATIONAL TESTING SERVICE, AN APPROACH FOR IDENTIFYING AND MINIMIZING BIAS IN STANDARDIZED TESTS 5 (1980) ("Given any paper and pencil test, chosen at random, and any particular minority group, it is well documented that as a group minorities can be expected to score lower."); Beth Dawson et al., *Performance on the National Board of Medical Examiners Part I Examination by Men and Women of Different Race and Ethnicity*, JAMA, Sept. 7, 1994, at 674, 674 (finding that men scored higher than women and that whites scored higher than other racial groups on the multiple choice exam administered to all prospective medical doctors).

See also the following cases: *Sharif v. New York State Educ. Dep't*, 709 F. Supp. 345 (S.D.N.Y. 1989) (holding that use of SAT adversely affected female applicants for New York State merit scholarships); *Larry P. v. Riles*, 495 F. Supp. 926 (N.D. Cal. 1979) (noting persistent disparate impact of intelligence tests on blacks and noting existence of cultural bias), *aff'd in part and rev'd in part*, 793 F.2d 969 (9th Cir. 1984).

Some, although certainly not all, of the exclusionary impact of tests and informal assessment on women and people of color is a result of their lower socio-economic positions.¹⁷⁰ Women and people of color are overrepresented in the lower socio-economic sectors. Indeed, because black middle-class status is so recent, income data tends to overstate the actual socio-economic level of blacks.¹⁷¹ This is particularly so because of the stark differences in accumulated assets between whites and blacks. On average, whites have nearly twelve times the median net worth of blacks; in terms of net financial assets, the average black household has no nest egg whatsoever.¹⁷² For these reasons, the exclusionary impact of existing selection practices on people of color and women makes visible a pattern of class-based exclusion that would otherwise remain hidden.

There is also evidence that certain tests are less predictive for some groups than others. For example, although women as a group perform less well than males on the SAT, they equal or outperform men in first-year college grade point average, the most common measure of successful performance.¹⁷³ Similar patterns have been detected in the results of the ACT and other standardized college selection tests.¹⁷⁴

Supplementing class rank with the SAT also decreases black acceptances and black admissions.¹⁷⁵ Yet, "when the admissions rate is controlled, the SAT cannot admit successful blacks or reject unsuccessful ones with substantively greater accuracy than high school rank."¹⁷⁶

170. The College Board uses this correlation between class and gender to help explain the differential performance of women and men on the SAT. See CAMERON, *supra* note 42, at 15 ("Women are more likely to come from a family where neither parent attended college. Family income is found to be higher for men as well. Differences in population size, academic background, and socioeconomic status help explain the difference in mean scores.").

171. See MELVIN L. OLIVER & THOMAS M. SHAPIRO, *BLACK WEALTH/WHITE WEALTH: A NEW PERSPECTIVE ON RACIAL INEQUALITY* 5-6 (1995):

[T]he best indicator of the sedimentation of racial inequality is wealth. . . . White and black incomes are nearing equality for married-couple families in which both husband and wife work: in 1984 such black households earned seventy-seven cents for every dollar taken home by their white counterparts. Yet in 1984 dual income black households possessed only nineteen cents of mean financial assets for every dollar their white counterparts owned. A black-to-white ratio of 77 percent represents advancement and is cause for celebration, while a 19 percent wealth ratio signals the persistence of massive inequality.

Id. (emphasis added).

172. *Id.* at 86 ("Whites possess nearly twelve times as much median net worth as blacks, or \$43,800 versus \$3,700. In an even starker contrast, perhaps, the average white household controls \$6,999 in net financial assets while the average black household has no net financial assets or nest egg whatsoever.").

173. ROSSER, *supra* note 169, at 4.

174. *ETS Developing "New" GRE*, *supra* note 106, at 11 ("Research . . . shows the GRE underpredicts the success of minority students. And an ETS Study concluded the GRE particularly underpredicts for women over 25, who represent more than half of female test-takers.").

175. CROUSE & TRUSHEIM, *supra* note 43, at 103 (finding that black admissions decrease from 74.6% to 57.7% when high school rank is supplemented with SAT scores).

176. *Id.* at 107. Crouse and Trusheim discuss a study that used statistical methods to simulate the projected performance of rejected applicants to college. It showed that the SAT reduces the number

Studies show that the group of black applicants rejected based on their SAT scores includes both those who would likely have failed and those who would likely have succeeded, and that these groups offset each other. Consequently, the rejection of more blacks as a result of using SAT scores "does not translate into improved admissions outcomes. The SAT does not improve colleges' ability to admit successful blacks and reject potentially unsuccessful ones."¹⁷⁷

Research has also demonstrated the discriminatory practices produced by unstructured subjective screening practices. Studies have shown that women and people of color are evaluated more negatively than white men with identical credentials¹⁷⁸ and are less likely to receive helpful mentoring.¹⁷⁹ Moreover, although personal connections and word-of-mouth recruitment continue to prevail as the most widely used methods of recruitment and selection,¹⁸⁰ women and people of color frequently remain outside the networks that lead to desirable jobs.¹⁸¹

of black false positives (applicants who would otherwise be admitted but would not be successful). But the SAT also increases the number of black false negatives (applicants who are rejected yet would be successful). *Id.* at 103-05. The results of this study suggest that

color blind admissions with the SAT increase colleges' black true negative [applicants who are rejected and who would be unsuccessful] admissions decisions and decrease their false positive decisions. Both results seem desirable. However, colleges will also decrease their true positive [applicants who are admitted and who will be successful] admissions decisions and increase their false negative decisions. Both these results seem undesirable.

Id. at 106.

177. *Id.* at 107-08.

178. See, e.g., BLUEPRINT FOR GENDER EQUITY, *supra* note 128, at 29 (finding that the absence of safeguards and validation permits manipulation of testing and interviews to disadvantage women). Surveys involving various professions find that the same work or the same resume is rated lower if attributed to a woman rather than a man. See Madeline E. Heilman & Melanie H. Stopeck, *Being Attractive, Advantage or Disadvantage? Performance-Based Evaluations and Recommended Personnel Actions as a Function of Appearance, Sex, and Job Type*, 35 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 202 (1985) (arguing that stereotypic characterizations mediate sex discrimination in personnel decision making); Barbara Reskin, *Bringing the Men Back In: Sex Differentiation and the Devaluation of Women's Work*, in THE SOCIAL CONSTRUCTION OF GENDER, 141, 145-46 (Judith Lorber & Susan A. Farrell eds., 1991); Michele A. Paludi & William D. Bauer, *Goldberg Revisited: What's in an Author's Name*, 9 SEX ROLES 387 (1983). For a discussion of the dynamics of stereotyping in evaluation, see Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317 (1987); Mary F. Radford, *Sex Stereotyping and the Promotion of Women to Positions of Power*, 41 HASTINGS L.J. 471 (1990).

179. Guinier et al., *supra* note 68, at 73-75.

180. EZORSKY, *supra* note 31, at 15 ("Kathleen Parker of the National Center for Career Strategies was reported in 1990 as stating that over 80 percent of executives find their jobs through networking and that about 86 percent of available jobs do not appear in the classified advertisements.").

181. *Id.* at 15-16; Edward W. Jones, Jr., *Black Managers: The Dream Deferred*, in DIFFERENCES THAT WORK: ORGANIZATIONAL EXCELLENCE THROUGH DIVERSITY, *supra* note 145, at 65, 74-75 ("To get ahead, a person depends on informal networks of cooperative relationships. . . . Black managers feel they are treated as outsiders, and because of the distance that race produces they don't receive the benefit of these networks and relationships."); Wilkins & Gulati, *supra* note 90, at 558 ("[B]lack on average have less access to influential contacts and other informal networks that allow some other candidates to bypass the formal screening requirements.").

There is compelling evidence that "word of mouth recruitment where the existing labor pool is predominantly white male reduces the chances of women or people of color applying for the jobs, as do unions that influence or control hiring in well-paid jobs in construction, transportation, and printing when they recruit through personal contacts."¹⁸²

Children of alumni, who are overwhelmingly white, constitute between twelve and twenty-five percent of some of the top schools in the country.¹⁸³ Nepotism, networking, and word-of-mouth recruitment for positions in government and business advantage the children of those who occupy positions of influence within the system.¹⁸⁴ Legacy admissions, alumni preferences, the old-boy network, and numerous other departures from so-called objective merit standards favor white males and individuals from higher socio-economic backgrounds.¹⁸⁵ This is particularly apparent in higher education, where legacy admissions policies favor the children of highly educated and wealthy alumni, a disproportionate percentage of whom are white males.¹⁸⁶ These patterns of informal selection, which disadvantage women and people of color and privilege the "haves," may help explain data showing that white men

182. Harris & Narayan, *supra* note 7, at 20 (citation omitted).

183. Foster, *supra* note 8, at 143.

184. About one in five students at schools such as Harvard and Yale is a child of an alumnus. Mark Megalli, *So Your Dad Went to Harvard: Now What About the Lower Board Scores of White Legacies?*, J. BLACKS HIGHER EDUC., Spring 1995, at 71, 72 (finding that despite weaker GPAs, extracurricular activities, and SAT scores, legacy applicants enjoy twice as great a chance of being admitted to Harvard and Yale and almost three times as great a chance of being admitted to Princeton). A U.S. Department of Education Office of Civil Rights investigation of Harvard found that legacy preferences disproportionately help white applicants, because 96% of all living Ivy League alumni are white. *Id.* at 72.

185. See Ralph Frammolino et al., *UCLA Eased Entry Rules for the Rich, Well-Connected*, L.A. TIMES, Mar. 21, 1996, at A1 (reporting that chancellor and top aides gave back-channel admissions help to friends or relatives of donors, and that 75 privileged students were admitted while hundreds of others with better grades and higher SAT scores were turned away).

186. Lind, *supra* note 19, at 582-83 ("In an industrial, bureaucratic society in which access to wealth and power depend on educational credentials, alumni preference in university admissions is the managerial-professional equivalent of primogeniture. Legacy preference is affirmative action for the Haves."). According to one study, "the combined SAT scores of those students admitted under [such policies] are thirty five points below those of other students." Foster, *supra* note 8, at 143; see also John D. Lamb, *The Real Affirmative Action Babies: Legacy Preferences at Harvard and Yale*, 26 COLUM. J.L. & SOC. PROBS. 491 (1993) (finding that schools such as Harvard and Yale have been awarding extra points to alumni children since at least the 1920s, when schools sought ways to limit the burgeoning numbers of Jewish and Catholic students on campus); Weber, *supra* note 71 (finding that at Harvard fewer than 16% of regular applicants but close to 40% of legacy applicants are admitted); Frank Rich, *Journal: Class of '71*, N.Y. TIMES, June 5, 1996, at A21 (noting that Harvard's class of 1971 was 93% white and 77% male).

are overrepresented in top positions of management, *regardless of educational level*.¹⁸⁷

Thus, it is incontestable that the existing meritocracy excludes people based on their race, gender, and class status. It is also without question that the construction of our conventional meritocracy disproportionately includes people who are wealthy, male, and white. The existing meritocracy creates a modern-day aristocracy that gives further advantages to the already advantaged, and creates barriers for those who are not.¹⁸⁸

The question remains: is this highly unequal outcome fair? Based on at least some definitions of fairness, this exclusionary outcome may be justified, at least with respect to selection criteria that are not facially biased. Even if the "meritocracy" screens out women, people of color, and those of lower socio-economic status, it could be argued that those screens are fair if they serve an important function. The structure of this argument tracks the structure of disparate impact theory under Title VII: selection criteria that disproportionately exclude members of protected classes are not discriminatory if employers justify them as "job related for the position in question and consistent with business necessity."¹⁸⁹ If our existing selection standards both identify those who can perform best in the positions we seek to fill, and assess capacity to perform across axes of group difference (race, class, gender), then they are fair even if they are exclusionary of some groups and privileging of others.¹⁹⁰

187. EXECUTIVE, ADMINISTRATIVE, AND MANAGERIAL OCCUPATIONS BY
RACE, EDUCATIONAL ATTAINMENT, AND SEX, 1990

Degree	Percent in Occupations			
	White		Black	
	M	F	M	F
Less than a high school diploma	63.5	28.9	4.6	3.0
4 years of high school	49.5	46.1	1.9	2.5
1 to 3 years of college	55.2	38.8	2.7	3.2
4 or more years of college	67.5	26.6	2.9	3.0

TOTAL EXECUTIVES, ADMINISTRATORS, AND MANAGERS (Blacks and Whites only)				
	59.8	34.7	2.6	2.9

FEDERAL GLASS CEILING COMMISSION, GOOD FOR BUSINESS: MAKING FULL USE OF THE NATION'S HUMAN CAPITAL 75 (1995).

188. See Lind, *supra* note 19, at 579-86 (describing the "monopolization of the best positions in the managerial-professional elite by a hereditary or quasi-hereditary social class" through rules of professional licensing and education that favor this elite, and calling for a war on oligarchy that rewrites the "rules of the educational game, as well as the rules of professional accreditation . . . to make social mobility easier in America").

189. 42 U.S.C. §2000e-2(k)(1)(A)(i) (1996).

190. This assessment does not address the fairness concerns linked to distributive justice and participation. See *infra* Part IV.

We have shown, however, that the testocracy fails even on this measure; it does not reliably distinguish successful future performers from unsuccessful ones, even when supplemented by additional subjective criteria. Therefore, racial, gender, and socio-economic exclusion cannot legitimately be justified in the name of a flawed system of selection.

III

THE NEED FOR A NEW PARADIGM FOR SELECTION AND INCLUSION

We have seen how the stock affirmative action narrative normalizes and legitimates selection practices that are neither fair nor functional. Such scrutiny of the selection standards themselves rarely surfaces in judicial opinions about affirmative action. For example, in *Hopwood v. Texas*,¹⁹¹ the Fifth Circuit never questioned either the validity or fairness of the underlying criteria used to evaluate candidates for admission. The court assumed without discussion that students who scored higher on the Texas Index were necessarily more qualified for admission, and that distinctions based on relatively narrow differences in test scores constituted a legitimate basis for differentiating among applicants.¹⁹²

Similarly, in *Johnson v. Transportation Agency of Santa Clara County*,¹⁹³ the Supreme Court credited the legitimacy of a ranking system based on an interview process that appeared to have little accountability and was tainted by bias. Statistically insignificant two-point differentials were treated as important in establishing the superior qualifications of the plaintiff challenging the affirmative action system.¹⁹⁴ Sex could legitimately be treated as a "plus" factor to alter the results of the normal, albeit flawed, selection process.¹⁹⁵

Many others concerned about racial and gender justice have challenged the fairness and validity of conventional approaches to merit.¹⁹⁶ These critiques of merit are often linked to arguments

191. 78 F.3d 932 (5th Cir.), *cert. denied*, 116 S. Ct. 2582 (1996).

192. *Id.* This blanket acceptance of the underlying standards persisted in the face of evidence that the Texas Index was less predictive for black applicants than for others, Declaration of Martin M. Shapiro, *supra* note 69, and that "applicants selected for admission come from a relatively narrow band within the full range of scores, and a difference of few points does not necessarily correlate with more successful work in law school," *Hopwood v. Texas*, 861 F. Supp. 551, 563 (W.D. Tex. 1994) (referring to expert affidavits).

193. 480 U.S. 616 (1987).

194. *See Selmi, supra* note 8, at 1252-53.

195. *Id.* (noting that "no one ever questioned, or even mentioned" the assumption that Mr. Johnson was better qualified than Ms. Joyce).

196. *See, e.g., EZORSKY, supra* note 31, at 88-93; YOUNG, *supra* note 8, at 203-05; Fallon, *supra* note 33, at 815; Foster, *supra* note 8, at 157; Harris & Narayan, *supra* note 7, at 20-24; Kelman, *supra* note 8, *passim*.

justifying the departure from those conventional approaches to merit.¹⁹⁷ For example, Luke Harris and Una Narayan, who are eloquent in both their critique of the rhetoric of preferential treatment and their articulation of a vision of full citizenship as a justification for affirmative action, remain within the existing paradigm nonetheless.¹⁹⁸ They call for a shift in the moral and legal justification for affirmative action from preferential treatment to "equality and full citizenship as they relate to the rights of Blacks and other marginalized Americans."¹⁹⁹ Yet, the programs themselves remain as add-ons to existing, admittedly dysfunctional selection standards. In effect, this approach subscribes to the sameness-as-fairness view.²⁰⁰ Affirmative action is necessary to level the playing field.

This Section argues the importance of moving from affirmative action as an add-on to affirmative action as an occasion to rethink the organizing framework for selection generally.

A. *Plus Factors and Preserving the Status Quo*

Many of the justifications for relying exclusively or predominantly on race and gender plus factors to reach numerical hiring goals make short-term sense. These factors are easily measurable, and thus easy to monitor.²⁰¹ They produce quick results that are visible and concrete.²⁰²

197. See Selmi, *supra* note 8, at 1277-79 (linking critique of test score ranking to the legitimization of affirmative action programs that use racial or gender preferences).

198. See Harris & Narayan, *supra* note 7. Harris and Narayan take issue with those proponents of affirmative action who support the compensation rationale, under which "affirmative action is seen as offering preferences to women or members of racial minorities as reparation or compensation for past injustices." *Id.* at 14. These arguments focus too much attention on past individual injury, rather than patterns of continuing institutional exclusion or marginalization.

199. *Id.* at 4.

200. It implicitly accepts the continued application of merit standards, but employs a version of the sameness-as-fairness approach to justify limited departures from those standards. Because employers and admissions programs routinely depart from merit standards for alumni, friends, athletes, and others, it is fair to depart from those standards to promote racial and gender inclusion.

201. Cf. Local 28, Sheet Metal Workers Int'l Assoc. v. EEOC, 478 U.S. 421, 448 (1986) (justifying court-ordered affirmative action as necessary to monitor cases where employers are intransigent); Susan P. Sturm, *A Normative Theory of Public Law Remedies*, 79 GEO. L.J. 1355, 1360-67 (1991) (describing role of specificity in facilitating monitoring of remedies).

202. Several recent articles argue that affirmative action in the form of racial plus factors contributes to the overall efficiency of institutions. For example, Ian Ayres and Peter Cramton argue in a forthcoming article that racial and gender bidding preferences created extra competition and induced non-subsidized firms to bid higher in an auction for F.C.C. radio licenses. Ayres and Cramton show that, contrary to the criticism of the affirmative action program as a public giveaway, subsidizing women and people of color increased the government revenue by more than 15 percent. Ian Ayres & Peter Cramton, *Pursuing Deficit Reduction Through Diversity: How Affirmative Action at the FCC Increased Auction Competition*, STAN. L. REV. (forthcoming 1996) (manuscript at 2, 3-4, on file with authors). Thus, Ayres & Cramton conclude, affirmative action not only increases diversity, but may force non-preferred applicants to increase the quality of their bids.

Michael Selmi has also argued recently that "racial and gender preferences [may serve] as a positive measure to send signals to potential employees in order to attract individuals who might

Studies suggest that they have increased the participation of women and people of color in institutions that have adopted affirmative action plans.²⁰³ They avoid the difficult challenge of addressing the structural barriers to exclusion and to eliminating the exercise of bias in the process.²⁰⁴ In institutions with a pattern of excluding women and people of color, the hammer of measurable goals may be the only viable way to begin the process of inclusion.

The problem is that race- or gender-based departures from prevailing selection standards are not perceived, treated, or responded to in the same ways as are other departures from these standards. Race and gender characteristics are highly visible, unlike many of the categories of identity that also trigger departures from stated criteria and processes of selection. In an environment of economic uncertainty, affirmative action tends to be seen by vulnerable workers as an island of unfair treatment, rather than as the tip of the iceberg of departures from stated selection criteria.²⁰⁵ For example, Philadelphia ironworkers and Chicago firefighters who are not doing as well as their parents apparently focus their anger on black workers, whom they perceive as affirmative action beneficiaries, rather than on general economic conditions or on politicians or administrators making cutbacks.²⁰⁶

In addition, employers and universities often publicly announce affirmative action programs and justify deviations from their "normal" practices in terms of affirmative action. At the same time, they often fail to inform employees and students that considerations of merit strongly influenced the decision.²⁰⁷ So, for example, in a Law School seminar

otherwise have been discouraged by the presence or perception of employment discrimination." Selmi, *supra* note 8, at 1299. He further argues that affirmative action may have a positive impact on the effort level of all employees by providing "the previously missing perceptions of workplace opportunity that are necessary to stimulate worker effort," and by "increas[ing] competition which, in turn, increases effort and productivity." *Id.* at 1305-06. *But see* Ayres and Cramton, *supra*, at 64 (noting information risks in using affirmative action policies to increase firm profits).

203. See Jonathan S. Leonard, *What Promises are Worth: The Impact of Affirmative Action Goals*, 20 J. HUM. RESOURCES 1, 18 (1985) (finding that goals set in negotiations with the Office of Federal Contract Compliance "have a measurable and significant correlation with improvements in the employment of minorities and females at reviewed establishments"); *see also* Alison M. Konrad & Frank Linnehan, *Formalized HRM Structures: Coordinating Equal Employment Opportunity or Concealing Organizational Practices?*, 38 ACAD. MGMT. J. 787, 808 (1995) (noting that equal-employment-opportunity and affirmative action efforts led to development of identity-conscious procedures, which were positively associated with employment status of protected groups).

204. See Selmi, *supra* note 8, at 1296-97 (suggesting that affirmative action might be a more efficient monitoring device than either searching for unbiased managers or establishing review processes).

205. This is particularly true in an election year, when affirmative action is used to deflect attention from underlying concerns about economic security. Consider, for example, the campaign launched by Senator Jesse Helms against his challenger, Harvey Gant. See *supra* note 23.

206. St. George, *supra* note 2, at A1; Uchitelle, *supra* note 13, at A1.

207. Jennifer L. Eberhardt & Susan T. Fiske, *Affirmative Action in Theory and Practice: Issues of Power, Ambiguity, and Gender Versus Race*, 15 BASIC & APPLIED SOC. PSYCHOL. 201, 215 (1994)

that we co-teach, the students of color almost unanimously reported that their classmates assumed that they were accepted to Penn because of affirmative action. In contrast, no one in the class knew or even surmised that a white male student was accepted by virtue of his status as the son of an alumnus.²⁰⁸

This heightened visibility of race- and gender-driven exceptions to "objectively derived" test-score-based selection plays into existing biases and stereotypes, particularly about race.²⁰⁹ Studies and public opinion polls show that dominant group members often evaluate women and people of color more harshly than they do dominant group members, and that they continue to hold stereotypes that reinforce the perception that people of color are less qualified and thus would not be selected if merit standards were used.²¹⁰ These views are particularly strong in situations where prevailing stereotypes tend to be more simplistic and linked to concerns about performance.²¹¹ Thus, visibility, combined with underlying racism or sexism, undermines the capacity of non-beneficiaries of affirmative action to see race- or gender-linked departures from selection criteria as fair.

Departures from "merit" standards that benefit dominant groups are often perceived as legitimate and related in some way to the interests of the institution. Many dominant group members see as legitimate in-

("Because the conditions surrounding a position obtained through affirmative action are often left ambiguous, nontargeted individuals may give their targeted colleagues little credit for the qualifications and talent they bring to the job."). This reliance on affirmative action to justify decisions not to hire, promote, or admit white men profoundly and negatively influences reactions to affirmative action. Clinical experiments demonstrate that when employers do inform employees that decisions were based on a combination of merit and diversity—the overwhelmingly common scenario—non-beneficiaries react much more favorably to affirmative action programs. *Id.* at 215-16; see also Kadaba, *supra* note 12, at A2 (finding that whites with firms with affirmative action programs were more supportive of certain race-targeted remedies and were more likely to believe that discrimination exists).

208. Videotape of Race Talk: Collaboration Through Conversation, Seminar on Critical Perspectives on the Law: Issues of Race and Gender, held at the University of Pennsylvania (Nov. 29, 1995) (on file with authors).

209. Eberhardt & Fiske, *supra* note 207, at 204-05:

[S]implistic and stereotypic judgments are made much more often about members of powerless groups. . . . In the same vein, recipients of affirmative action are thought to be in that position (i.e., powerless) because they do not have the talent, skills, or drive necessary to become members of the controlling, dominant group.

210. *Id.*

211. *Id.* Research suggests that gender stereotypes tend to be more complex and prescriptive than racial stereotypes. *Id.* at 208. The simplistic and descriptive character of racial stereotypes heightens the tendency to generalize about individual blacks, and may also heighten negative reactions to affirmative action benefitting black workers or students:

Blacks may be seen as a relatively homogeneous out-group whose members are largely unintelligent, unskilled, and therefore, undeserving of group-based rewards. Because the stereotypes are more simplistic, they offer fewer possibilities for perceiving affirmative action in a non-threatening manner.

Id.

stitutions' decisions to favor alumni as a strategy for fundraising and promoting institutional loyalty.²¹² In contrast, conservative political activists, with the help of the Supreme Court, have been very effective in delegitimizing reliance on race or gender as legitimate categories of consideration.²¹³ Affirmative action-based departures often appear to pit diversity concerns against concerns of merit and institutional efficiency.²¹⁴ Affirmative action as an add-on deflects attention from the role of diversity in enhancing the productive capacity and efficiency of organizations.

The perception of illegitimacy accompanying race- or gender-tinted assessments sometimes spills over to performance.²¹⁵ Claude Steele's recent studies suggest that self-consciousness of group performance negatively affects minority group members' performance on tests.²¹⁶ Existing approaches to merit selection frequently marginalize or disguise the contributions and talents of those with styles or methods that depart from the norm. These standards cannot be challenged solely from the perspective of the margins, because the challenges themselves then become marginalized. As in the electoral context, race is rein-

212. Fallon, *supra* note 33, at 869-76 (discussing values of cohesion and familiarity as merit characteristics); Foster, *supra* note 8, at 143 (describing tolerance of dramatic departures from prevailing standards of merit for alumni and athletes).

213. See *Adarand Constructors, Inc. v. Peña*, 115 S. Ct. 2097, 2119 (1995) (Scalia, J., concurring) ("To pursue the concept of racial entitlement—even for the most admirable and benign of purposes—is to reinforce and preserve for future mischief the way of thinking that produced race slavery, race privilege and race hatred."); *Miller v. Johnson*, 115 S. Ct. 2475, 2482-83 (1995) (delegitimizing acknowledgment of racial difference and insisting on color blindness across the board).

214. In addition, dominant group members are more likely to perceive prevailing standards of selection as basically fair, valid, and neutral. Much of the critique of the existing merit standards rests on an understanding of the culturally biased character of those standards. This conception of fairness differs from the conception often embraced by majority group members, who frequently emphasize process conceptions of fairness. They tend to view existing selection criteria as natural, inevitable, and basically valid ways of choosing applicants for positions. As long as everyone is treated the same, i.e., evaluated using the same process and standards, then the selection process is fair. This concept of fairness does not easily accommodate departures from the norm premised on an understanding of substantive inequality and bias. See Eberhardt & Fiske, *supra* note 207, at 205.

215. Studies show that powerful group members are less aware than powerless group members of the ways that their achievements are group-linked. White women who are selected for a position because of their gender are likely to have more negative self-evaluations than those selected by merit. There is no similar effect for men. Whereas white women selected because of their gender choose safer, easier tasks to perform, white men exhibit no difference in task choice as a function of selection procedure. *Id.* But cf. Hochschild, *supra* note 146 (suggesting that on balance rumors of inferiority are somewhat exaggerated).

216. Steele & Aronson, *supra* note 92 (attributing black underachievement to the possibility of being judged and treated stereotypically). Steele and Aronson conducted experiments contrasting the performance of black and white students on tests characterized as either diagnostic or not diagnostic of intellectual ability. They found that the diagnostic tests were more likely to activate the stereotype threat, and that test diagnosticity impaired the accuracy and rate of black participants' work. *Id.* at 808.

forced and reproduced as marginal within institutions, which remain basically unchanged.²¹⁷

Thus, the same culture of dominance that necessitates affirmative action undermines the capacity of rhetorical strategies to justify race- or gender-linked departures from standards perceived as normal and fair. Constructions of race hide the unfairness of the whole system. The strategy of using the critique of merit to justify supplementing or departing from otherwise-operative standards also undermines the power and legitimacy of the merit critique itself. It seems to embrace an approach to selection that it acknowledges is at best imperfect and at worst completely arbitrary.

In some respects, affirmative action programs have indirectly prompted some general changes in selection practices. For example, in response to evidence that the old-boy-network approach to recruitment excluded people of color and women, many firms and schools have abandoned informal networking in favor of advertising, posting, and active searches, which expand the applicant pool.²¹⁸

However, most affirmative action programs in place do not respond to the bias and invalidity of selection practices by posing a direct and systemic challenge to those practices. Instead, they attempt to compensate for those inadequacies as they affect women and people of color by roughly approximating an outcome that might be achieved in a fair, unbiased world. They do not offer an alternative approach to defining or identifying qualified candidates. They have not suggested new ways of determining whether affirmative action programs have in fact leveled the playing field, as opposed to creating a new and more favorable set of rules for their beneficiaries. Perhaps for this reason, commentators have begun to emphasize the importance of retention, rather than hiring, as a serious challenge to inclusion and diversity.²¹⁹

This approach maintains the dichotomy between diversity and merit. Notwithstanding their bias and inadequacy, existing merit selec-

217. See Lani Guinier, *No Two Seats: The Elusive Quest for Political Equality*, 77 VA. L. REV. 1413, 1443-47 (1991); Lani Guinier, *The Triumph of Tokenism: The Voting Rights Act and the Theory of Black Electoral Success*, 89 MICH. L. REV. 1077, 1112-28 (1991) [hereinafter Guinier, *The Triumph of Tokenism*] (arguing that black electoral success from majority black districts may create electoral opportunities for black advocates who are then marginalized within the legislature and citing studies of small group interaction showing that minority views often are ignored, especially in competitive decision making).

218. See Edward J. Giblin & Oscar A. Ornati, *Beyond Compliance: EEO and the Dynamics of Organizational Change*, PERSONNEL, Sept.-Oct 1975, at 38 (describing the rationalization of human resource systems after affirmative action programs are implemented).

219. See Thomas, *supra* note 145, at 28; *Feds Fire Blacks More Often Than Whites*, CAPITAL, Apr. 1, 1995, at A5 (reporting that a study commissioned by the Office of Personnel Management found that black federal employees were fired at nearly twice the rate of their white, Hispanic, or Asian counterparts, controlling for a score based on other factors, including performance ratings, seniority, and education); see also De Witt, *supra* note 141, at A19 (noting same).

tion standards remain unchanged, and affirmative action hires are justified through considerations of diversity coupled with fairness challenges to existing standards. As we will discuss below, most institutions lack reliable ways of assessing performance, so that decision makers often fall back on the same unfair and invalid standards to determine the performance of those hired through affirmative action. This approach leaves a vacuum of focus or justification for selection decisions under the current system. This vacuum can create the impression that critics of existing merit systems reject the idea of merit itself—the idea that it is legitimate for schools, employers, and businesses to seek ways of selecting people who are capable of performing. By failing to apply the merit critique to the selection process itself, progressives also fail to identify as central to the affirmative action project the need to articulate new, more valid and inclusive approaches to selection norms and practices generally.

B. The False Promise of Prediction

To succeed in the search for fairness and merit in selection, it is crucial to move beyond the idea of one-size-fits-all testing. We are not suggesting that the solution is to develop a new, less biased, equally universal test that more accurately predicts future performance. Instead, we are challenging the idea of prediction as the organizing framework for selection.

The standard approach proceeds essentially as if selection were a matching process fine-tuned to apply predetermined criteria of performance to applicants and rank order the results. This assumes that the capacity to perform—functional merit—exists in people apart from their opportunity to work on the job; in other words, it assumes that people have the qualities necessary to do the job independently of their opportunity to do the job. It further assumes that institutions know in advance what they are looking for, and that these functions will remain constant across a wide range of work sites and over time. In particular, standard approaches to testing and test validation assume that we can predict what the job will require in the future, based on how it has operated in the past.²²⁰

In fact, neither the candidates for positions nor the positions themselves remain fixed. Actual performance often correlates best with on-

220. This assumption is built into the methodology of test validation. One of the critical steps of developing a test is to perform a job analysis, which proceeds by interviewing workers about the most significant aspects of their work. This job analysis then determines the functions and qualities that are necessary for future occupants of the position. See *Gillespie v. Wisconsin*, 771 F.2d 1035, 1039-40 (7th Cir. 1985) (describing the process of conducting a job analysis), *cert. denied*, 474 U.S. 1083 (1986); Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607.14(A)-(B) (1996) (setting out standards for job analysis and its relationship to test validation).

the-job training. Those people who do well learn their job on the job. Often those people who have been given an opportunity to *do* a job perform because they have been given an opportunity to *learn* the job. It is the opportunity to learn a job, a craft, or a skill that often predicts successful on-the-job performance.²²¹ This phenomenon tracks the way many experts "learn" their expertise. Experts become skilled as a result of the opportunity to develop their expertise by tackling actual problems.²²²

On-the-job learning has assumed even greater significance in the context of today's rapidly changing economy.²²³ Recent studies suggest that employers have become increasingly skeptical about the capacity of educational institutions to prepare applicants to meet the demands of the workplace, and have begun to build continual retraining into the job description.²²⁴ Unstable markets, technological advances, and shorter product cycles have created pressures for businesses to increase the flexibility and problem-solving capacity of workers.²²⁵ These pressures in turn increase the significance of on-the-job training and continuing education in enabling workers to develop the skills necessary to per-

221. For example, on the occasion of Robert MacNeil's retirement from the MacNeil/Lehrer News Hour on PBS, Jim Lehrer gave an interview on NPR. He was asked whether he ever went back to listen to the early tapes of the broadcast. Lehrer said, not very often; he was too embarrassed. He didn't know what he was doing back then. He followed a scripted seven-question format, and did not listen to and engage the subject. He learned how to conduct a good interview over time. He learned on the job. He was given the opportunity to learn his job. Interview with Jim Lehrer, Co-host, *McNeil/Lehrer News Hour*, (National Public Radio, Oct. 20, 1995).

A former labor lawyer, now a professor of law at Georgetown, explained this phenomenon based on his experience defending employment discrimination lawsuits. He said that those given the chance to succeed are in fact most likely to succeed. What we mean by merit is often the same thing as opportunity: those who are given the opportunity to go to Harvard do better in life than those who are denied that opportunity. See Conversation Between Lani Guinier and Michael H. Gottesman, Professor of Law, Georgetown Law School, former partner, Bradhoff & Kaiser, Washington, D.C. (Feb. 1993).

222. See Daniel A. Farber & Suzanna Sherry, *Telling Stories Out of School: An Essay on Legal Narratives*, 45 STAN. L. REV. 807, 821 (1993) (citing studies of expert decision making to support conclusion that expertise does not consist simply of knowing more facts or rules, but that it "involves the skill of picking out the key features of a new situation," a skill that is "learned primarily through experience with large numbers of past situations").

223. COMMISSION ON THE FUTURE OF WORKER-MANAGEMENT RELATIONS, DEPARTMENT OF LABOR, FACT FINDING REPORT 6-10 (1994).

224. See, e.g., THOMAS A. KOCHAN & PAUL OSTERMAN, *THE MUTUAL GAINS ENTERPRISE* 29-35 (1994) (discussing increased need for skilled employees in the current economy and the inability of the educational system to meet this need, and citing examples of firms increasing their emphasis on training).

225. See, e.g., PETER DOERING ET AL., *TURBULENCE IN THE AMERICAN WORKFORCE* 4 (1991) (discussing importance of on-the-job learning due to redeployment and layoffs); Michael J. Piore & Charles F. Sabel, *THE SECOND INDUSTRIAL DIVIDE* 273 (1984) ("Production workers must be so broadly skilled that they can shift rapidly from one job to another; even more important, they must be able to collaborate with designers to solve the problems that inevitably arise in execution.").

form.²²⁶ Under these circumstances, access to these on-the-job training opportunities will determine functional merit—the capacity to perform the functions necessary to meet the goals of the enterprise.

The concept of selection as a matching process also presumes that institutions have a relatively clear idea of what they are looking for, what they value, and the relationship of particular jobs to those institutional goals. Even in a relatively stable economic and technological environment, there is little indication that institutions engaging in selection have attempted to articulate goals, much less developed a basis for measuring successful performance of those goals.²²⁷ The question of how to define successful performance, of both institutions and particular actors within them, is a critical step in developing fair and valid selection criteria and processes. Yet, it is one that is in its infancy in most institutional settings.

Defining successful performance has also become more complicated in the current economic and political environment. Traditional measures of success, such as short-term profitability or increases in marginal productivity, do not fully define successful performance, and may in fact distort the capacity to evaluate and monitor activities.²²⁸ In addition, standards must increasingly change to adapt to technological developments and changes in consumer demands.²²⁹ Scholars of economic organization and human resources now emphasize the impor-

226. Sara Rimer, *A Hometown Feels Less Like Home*, N.Y. TIMES, Mar. 6, 1996, at A1, A18 (reporting that more than 20% of the students in a Dayton community college already have undergraduate degrees, and are back in college to learn new job skills because "[y]ou have to keep growing dramatically if you're really going to keep a job").

227. See 2 JOHN M. BRION, ORGANIZATIONAL LEADERSHIP OF HUMAN RESOURCES: THE KNOWLEDGE AND THE SKILLS 736 (1989) ("[L]ittle is known by the managers about the job's requirements-beyond-the-technical, the nature of the individuals they need to fill them, or the characteristics of the functional department's long range [human resource] needs. Hiring and placement are done largely by intuition. . .").

228. See Charles F. Sabel, A Measure of Federalism: Assessing Manufacturing Technology Centers 9 (Feb. 22, 1995) (unpublished manuscript, on file with authors) [hereinafter Sabel, A Measure of Federalism] ("[N]o one performance measure or basket of such indicators is . . . a robust predictor of a firm's likelihood of success. . . . Only when firms are [on the verge of bankruptcy] does measurement of their performance yield a consistent result."); see also KOCHAN & OSTERMAN, *supra* note 224, at 19-43 (discussing measures of macroeconomic welfare), 45-77 (discussing need for new human resource policies emphasizing a broad range of employee skills), 169-90 (advocating expanded training programs to implement workplace changes); Charles F. Sabel, *Learning by Monitoring: The Institutions of Economic Development*, in THE HANDBOOK OF ECONOMIC SOCIOLOGY 137 (Neil J. Smelser & Richard Swedberg eds., 1994) [hereinafter Sabel, *Learning by Monitoring*] (arguing that standards for monitoring success should evolve as market demands change).

229. Cf. Sabel, *Learning by Monitoring*, *supra* note 228, at 144 ("The danger is that changed market conditions, and especially some innovation in process or product, so alter prevailing performance criteria that the original reference point becomes irrelevant."); Sabel, A Measure of Federalism, *supra* note 228, at 11 (reporting need for continuous evaluation and redefinition of indicators of performance).

tance of developing complex, interactive, and holistic approaches to measuring both institutional and individual performance.²³⁰ One-size-fits-all, matching approaches to selection do not easily accommodate this move toward more dynamic and interrelated assessments of successful performance.

Current selection approaches also focus on the decontextualized individual. Merit is something that an individual possesses in the abstract and demonstrates through a test or interview. But team performance may be as or more important in today's environment.²³¹ Increasingly, work requires the capacity to interact effectively with others, and the demands of the economy are moving in the direction of more interactive, team-oriented production.²³² The significance of integration of functions and collaboration with workers from various fields is growing.²³³ The capacity to adapt to rapid changes in technology, consumer preferences, and fluid markets requires greater collaboration at every level.²³⁴ Paper-and-pencil tests do not measure or predict an

230. John G. Belcher, *Gainsharing and Variable Pay: The State of the Art*, COMPENSATION & BENEFITS REV., May-June 1994, at 50, 51 (advocating the use of a family of measure approach, which "utilizes multiple, independent measures to quantify performance improvement"); Sabel, *A Measure of Federalism*, *supra* note 228, at 11 (urging discovery of "contextual character of performance indicators" and noting that firms develop a "'family of measures,' each element of which, on strategic reflection, plausibly creates incentives for a distinct aspect of improvement").

231. See GARDNER, *supra* note 21, at 172 (summarizing research suggesting that "successful performance of a task may depend upon a team of individuals, no single one of whom possesses all of the necessary expertise but all of whom, working together, are able to accomplish the task in a reliable way"); cf. Fullilove & Treisman, *supra* note 140, at 466-67 (reporting that students in study groups tend to perform much better in mathematics than those who study alone).

232. See PETER F. DRUCKER, *THE NEW REALITIES* 207-31 (1989) (discussing changes in work arrangements and management techniques necessitated by the growing importance of information); Mark Barenberg, *Democracy and Domination in the Law of Workplace Cooperation: From Bureaucratic to Flexible Production*, 94 COLUM. L. REV. 753, 879-903 (1994); Robert B. McKersie & Richard E. Walton, *Organizational Change*, in *THE CORPORATION OF THE 1990s* 244, 249-50, 255-56 (Michael S.S. Morton ed., 1991) (discussing greater interdependence among workers necessitated by technological change); Paul Osterman, *Impact of IT on Jobs and Skills*, in *THE CORPORATION OF THE 1990s*, *supra*, at 220, 220-43 (discussing effects of technological change on the nature of modern jobs and identifying greater overlap among worker responsibilities as one of the effects); John F. Rockart & James E. Short, *The Networked Organization and the Management of Interdependence*, in *THE CORPORATION OF THE 1990s*, *supra*, at 189 (discussing increased interdependence introduced by information technology, including shared goals, work, decision making, and responsibility).

233. Recent studies suggest that this is most true for technical and professional workers, but that collaboration, integration, and rotation of functions have also assumed increased significance for blue-collar and line workers. See COMMISSION ON THE FUTURE OF WORKER-MANAGEMENT RELATIONS, *supra* note 223, at 29-61 (discussing, among other workplace changes, "employee participation" programs, including team-based work structures, job rotation, and worker-management cooperation). See generally KOCHAN & OSTERMAN, *supra* note 224, at 29-37, 79-109.

234. Although there is debate about the degree of fundamental change in approaches to management, a significant portion of private businesses have adopted some form of collaborative or team-oriented production. EDWARD E. LAWLER III ET AL., *EMPLOYEE INVOLVEMENT AND TOTAL QUALITY MANAGEMENT: PRACTICES AND RESULTS IN FORTUNE 1000 COMPANIES* (1992) (analyzing the employee-involvement programs many corporations have adopted); Paul Osterman,

individual's capacity for creativity and collaboration. As we discuss below, collaboration also enhances success in creative problem-solving.²³⁵

Assessment through opportunity to perform works better than testing for performance. Various studies have shown that "experts often fail on 'formal' measures of their calculating or reasoning capacities but can be shown to exhibit precisely those same skills in the course of their ordinary work."²³⁶ Those who assess individuals in situations that more closely resemble actual working conditions make better predictions about those individuals' ultimate performance.²³⁷ Especially if those assessments are integrated into day-to-day work over a period of time, they have the potential to produce both better information about workers and better workers.²³⁸

Moreover, many of those who are given an opportunity to perform, even when their basic preparation is weaker, usually catch up if they are motivated to achieve. Indeed, a recent study of a twenty-five-year policy of open-admissions enrollment at the City University of New York found that the school was one of the largest sources in the United States of undergraduate students going on to earn doctorates, even though many of its undergraduates come from relatively poor backgrounds and take twice as long to complete their bachelor's degree.²³⁹

How Common is Workplace Transformation and Who Adopts It?, 47 INDUS. & LAB. REL. REV. 173, 176-78 (1994) (finding that over 50% of firms surveyed had introduced at least one innovation such as quality circles and work teams, and that 36.6% have at least two practices in place with 50% or more of employees involved in each).

The issue of the implications of these new forms of organizational governance for racial and gender inclusion and for legal regulation of the workplace in general is the subject of a forthcoming article by Susan Sturm entitled *Race, Gender and the Law in the Twenty-First Century Workplace*.

235. See *infra* Part IV.D.

236. GARDNER, *supra* note 21, at 172.

237. *Id.* at 175.

238. Cf. Elizabeth Badger, *Finding One's Voice: A Model for More Equitable Assessment* 8-14 (May 1996) (unpublished manuscript, on file with authors) (containing description of assessment that acknowledges each student's unique voice and encourages continuous evaluation of student work). See generally GARDNER, *supra* note 21, at 174-83 (describing the general features of a new approach to assessment).

239. See Arenson, *supra* note 159, at A1, B4 (citing study of open-admissions policy at City University of New York (CUNY) that found more than half of the students eventually graduated, even though it took many as long as ten years to do so). Many of these students had to work full time while they attended college. According to Professor David Lavin, one of the co-authors of the CUNY study, open admissions "provided opportunities that students used well, and that translated into direct benefits in the job market and clearly augmented the economic base." *Id.* Similarly, at Haverford College, professors of biology, chemistry, and mathematics found that many students of color with weak preparation in the natural sciences took two years to catch up with their better prepared peers. However, by junior year, those same students managed to excel, having overcome their initial disadvantages. Interview by Lani Guinier with 11 Haverford College Professors, in Haverford, Pa. (Apr. 4, 1996).

This phenomenon of students markedly improving in performance after receiving the opportunity to participate in exacting on-site training has also been observed in programs that provide incoming minority college freshmen rigorous preparation. For example, the grade point average of black and Latino freshmen who pre-enrolled in an intensive, five-week summer course exceeded the overall average in the Georgia Tech engineering program. "The performance gap between minority engineering students and white ones [was] eliminated."²⁴⁰

Thus, prediction as the model of selection has created an illusion of precision and validity that disables institutions from developing more dynamic and functional ways of choosing qualified candidates.²⁴¹ Prediction cannot substitute for the actual opportunity to perform as a basis for both equipping applicants to perform and determining whether they can in fact perform.

IV

RECLAIMING MERIT AND FAIRNESS: OPPORTUNITY AND ACCOUNTABILITY

We are seeking to locate affirmative action within a broader progressive agenda. It is our view that affirmative action has the potential to play a pivotal role in sparking a dialogue about how to reconceive an approach to selection that will benefit everyone.²⁴² We also believe that this reconceptualization of selection is a critical step in the pursuit of racial and gender justice.

We begin by asking the following questions: Is there a way to push the debate about affirmative action forward? Can we find examples that will point us in the direction of adapting to the changing work demands and demographics of the 21st century? How can we situate the conversation about racial and gender justice within a practice of fair, dynamic, and effective selection? We do not attempt here to provide a definitive

240. Ronald Smothers, *To Raise the Performance of Minorities, A College Increased Its Standards*, N.Y. TIMES, June 29, 1994, at A21.

241. See, e.g., Price, *supra* note 119, at 1, 4 (describing his own success in law school despite a weak showing on the law boards, and concluding that "[w]e need new techniques for identifying potential and nurturing talent[, a]nd we must give institutions and employers broad license to take calculated risks on candidates").

242. The goal is to bridge the real or imagined distance between people of color in urban cities and whites in suburban America, so that participants begin to move from individual positions to the collective interest. The goal is to get participants to understand and respect their differences of perspective, to remove barriers to mutual understanding, and to overcome the belief that our differences invariably and permanently separate us. Cf. Charisse Jones, *An Act of Youthful Savagery Stuns a Suburb*, N.Y. TIMES, Nov. 19, 1994, at A1, A9 (describing the shock and shame of suburban Abington residents at a teenager's violent death at the hands of fellow suburban teenagers: "'Maybe we can learn violence is not a city thing,' said Stephen Lutz, a 17-year-old senior. 'It's not a suburban thing. It's not an Abington thing. It's a society thing. It can touch you anywhere.'" *Id.* at A9.

answer to these questions, or a universal model for approaching them. Indeed, development of such a universal, top-down solution would likely replicate many of the limitations of the current approach. Instead, we hope to provoke a dialogue—a series of structured, ongoing, participatory, and locally grounded conversations that can begin to reorient our standards and processes of defining and pursuing merit.²⁴³

The dual concerns of racial and gender justice and economic revitalization challenge us to push toward a new stage of development in our approaches to merit selection. The first generation of practice relied heavily on informal, private, unstructured decision making. Many scholars have written about the dangers of informal, purely-process approaches to problem solving and dispute resolution.²⁴⁴ Informal, private dispute resolution replicates the power dynamics and exclusionary practices that frequently prompt a move to formal, rule-bound, adversarial decision making. At the same time, our analysis suggests that purely formal rule-bound approaches to selection will not meet the dual challenge of inclusiveness and economic revitalization.

The next step requires that we take account of the critiques of informality *and* formality. One promising direction builds on the strengths of both by encouraging collaboration, integration of diverse approaches, and rotation of power, while maintaining mechanisms to protect against abuses of power, to assure participation, to equalize influence over decision making, and to retain the capacity to pursue substantive goals of equality and productivity.²⁴⁵ We are not suggesting that this new direction is the only alternative. Nor do we argue that such an alternative can or should be imposed wholesale by courts on intransigent employers or schools, although the law could be used to create additional incentives for decision makers to move in this direction.²⁴⁶ In

243. Cf. Charles Sabel, *Bootstrapping Reform: Rebuilding Firms, the Welfare State, and Unions*, 23 POL. & SOC'Y 5-48 (1995) (urging a similar approach in the context of rebuilding regional economies and the labor movement).

244. See Richard L. Abel, *The Contradictions of Informal Justice*, in 1 *The Politics of Informal Justice* 267 (Richard L. Abel ed., 1982); Richard Delgado et al., *Fairness and Formality: Minimizing the Risk of Prejudice in Alternative Dispute Resolution*, 1985 WIS. L. REV. 1359 (arguing that deformalization may increase risk of class-based prejudice); Tina Grillo, *The Mediation Alternative: Process Dangers for Women*, 100 YALE L.J. 1545 (1991) (describing dangers to women in mandatory child custody mediation).

245. Cf. Sturm, *supra* note 201, at 1427-44 (articulating a model of structured deliberation in the context of public remedial decision making).

246. For a modest proposal moving in this general direction, see Tracy A. Baron, Comment, *Keeping Women Out of the Executive Suite: The Courts' Failure to Apply Title VII Scrutiny to Upper-Level Jobs*, 143 U. PA. L. REV. 267, 309 (1994) (suggesting that courts take into account the arbitrariness of the employer's decision-making system before deciding how much deference the employer's subjective judgment deserves); cf. IAN AYRES & JOHN BRAITHWAITE, *RESPONSIVE REGULATION: TRANSCENDING THE DEREGULATION DEBATE* 4 (1992) (advancing as central to responsive regulation "the idea that escalating forms of government intervention will reinforce and help constitute less intrusive and delegated forms of market regulation").

environments dominated by active and committed racists or sexists who have succeeded in keeping out women and people of color, participatory decision making will only reinforce an unfair status quo.

In the next Section, we sketch out one framework that admittedly depends for its implementation on active engagement by a critical mass of participants who are committed to long-term productivity and who recognize the significance of racial and gender inclusion in achieving this goal. We show how such an approach might integrate both selection and racial and gender inclusion into the overall decision making of an organization. We then discuss how this new model relates to the concepts that have been so central to the assault on affirmative action—fairness and merit. Finally, we introduce ideas about democratic opportunity as a normative basis for rethinking selection.

A. An Emerging Model of Selection through Experience and Structured, Participatory Assessment: Outlines and Stories

We are proposing a shift in the model of selection from prediction to performance. This model builds on the insight that the opportunity to participate creates the capacity to perform, and that actual performance offers the best evidence of capacity to perform. There simply is no substitute for experience, both in equipping people to perform and in producing informed judgments about the functional capacity of candidates. This approach shifts the emphasis away from the design of an instrument that is separate from the performance of the job, but that can be correlated with success in that job.²⁴⁷ Instead, the emphasis is on thinking creatively about how evaluation can proceed through the observation of applicants engaged in the work of those positions. The model also emphasizes the importance of creating opportunities to succeed and of structuring fair, inclusive, and participatory mechanisms to define and assess successful performance. This approach thus embeds performance and inclusion in the design of the selection process.²⁴⁸

This approach to selection sets up three critical challenges: (1) how to integrate the assessment process into the activities of the organization;

247. The employers and educational institutions that utilize standardized tests purportedly identify important elements of behavior central to the position under consideration, such as analytical ability; develop test instruments that purportedly measure those elements; and then attempt to correlate performance on the test with performance on the job. See Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607.14 (1996). As Part II, *supra*, demonstrated, most standardized tests fall short at each step of the analysis.

248. For similar approaches, see GARDNER, *supra* note 21, at 161-62, 173 (urging that new approaches to assessment draw on the virtues of the apprenticeship model); Badger, *supra* note 238, at 8-14 (urging importance of embedding assessment in the process of teaching and learning); see also David A. Thomas & Robin J. Ely, Managing Diversity for Organizational Effectiveness 3 (Oct. 21, 1994) (unpublished manuscript, on file with *California Law Review*) (critiquing the prevailing assumption that diversity should not affect the actual workings of the organization).

(2) how to integrate considerations of inclusion and diversity into the structure and process of decision making about selection; and (3) how to develop functional, performance-based thresholds for participation, and mechanisms of evaluation that are accountable to concerns of both performance and inclusion. For lack of a better term, we refer to this model of selection as "structured participatory decision making." By this, we mean decision making about merit, selection, promotion, and discharge that is flexible, public, accountable, and fully participatory.

This approach is closer to content-validation approaches to testing, which purport to create tests that closely approximate the tasks to be performed on the job.²⁴⁹ However, our approach differs in a critical respect. It does not create a new test or develop simulations of tasks that make up a small part of the overall job.²⁵⁰ Instead, its goal is to give applicants a meaningful opportunity to learn and to perform in the position as part of the application process. The challenge is as much one of organizational design as of assessment methodology. Can we think creatively about how to structure opportunities to learn and perform as part of the process of seeking new positions?

We are suggesting that the process of selection be reconceptualized from a static one, where measurement is detached from productivity, to a dynamic one, where feedback is integrated into productivity. At the level of individual performance assessment, this means looking less to one-shot, predictive tests and more to performance-based evaluation.

At the very least, decision makers would only take into account "true scores," meaning the bands or zones of scores that are reliably different from one another.²⁵¹ Tests would no longer serve to rank applicants in order of their scores. Indeed, testers need not even report the numerical score to the applicant or the decision maker. Instead, they would report the band score indicating statistically significant differentials in test performance. In addition, test scores would not be used as independent screens that function as prerequisites for further consideration. Instead, they would serve as one factor in the overall decision-making process. Both of these moves would decrease the centrality of the test scores in the decision-making process and force decision makers to develop other ways of distinguishing between candidates who are in-

249. Content validation is preferred by employers because it is often the only feasible way to validate a test. As the court noted in *Guardians Ass'n v. Civil Serv. Comm'n*, 630 F.2d 79, 92 (2d Cir. 1980), *cert. denied*, 452 U.S. 940 (1981): "This content-construct distinction has a significance beyond just selecting the proper technique for validating the exam; it frequently determines who wins the lawsuit. Content validation is generally feasible while construct validation is frequently impossible."

250. For a critique of the reliability of content validation for complex jobs, see Kelman, *supra* note 8, at 1210-12.

251. See Selmi, *supra* note 8, at 1275.

distinguishable based on their testing results. They would also undercut the legitimacy of a claim that an applicant necessarily deserves a particular job or position because he scored a few points higher on an objective test.

A more fundamental change resulting from our framework would be a shift away from reliance on tests as a means of distinguishing among candidates. One way to do this would be to use tests only as a floor to screen out those individuals who could not learn to perform competently with adequate training and mentoring. Another alternative would be to discontinue tests as a required part of the selection process.²⁵²

Of course, decreasing reliance on tests to rank candidates would create the need to develop other ways of distinguishing among applicants. There is no single, uniform solution to this problem. One approach would be a variation on the concept of a lottery: the distribution of opportunity to participate among relatively indistinguishable candidates by chance. Concerns about a lottery's insensitivity to particular institutional needs or values could be addressed by increasing the selection prospects of applicants with skills, abilities, or backgrounds that are particularly needed by the institution. A weighted lottery may indeed be the fairest and most functional approach for some institutions. Particularly in the education arena, where opportunity lies at the core of the institution's mission, a lottery may be an important advance.²⁵³ However, in many contexts a lottery may not be a viable option. Also, the lottery approach would not necessarily require an institution to engage in the process of defining its direction.

A more institutionally grounded approach might work in non-educational contexts. In some jobs, for example, decision makers would assume responsibility for constructing a dynamic and interactive process

252. See CROUSE & TRUSHEIM, *supra* note 43, at 148 ("[W]e think that both small and large selective colleges and universities that now use the SAT could drop the test with no serious negative consequences. Their admission rates for blacks and lower-income applicants would probably increase if they do not now use a quota."). A growing number of colleges has either abandoned the use of SAT/ACT scores in the selection process, or has made the submission of test scores optional. See FAIRTEST, 241 SCHOOLS WHERE SAT AND ACT SCORES ARE OPTIONAL FOR ADMISSION INTO BACHELOR DEGREE PROGRAMS (undated published sheet). There is also a growing movement in education to develop more performance-based, authentic methods of assessing students. See RUTH MITCHELL, TESTING FOR LEARNING: HOW NEW APPROACHES TO EVALUATION CAN IMPROVE AMERICAN SCHOOLS 20-21 (1992) (advocating move from testing to performance assessment); Badger, *supra* note 238, at 8-14 (proposing performance-based assessment model).

253. See *infra* Part IV.E (discussing the role of the institution in the democratic process); see also Peter E. Rosenfeld, *Inside the Meritocracy Machine*, N.Y. TIMES, May 19, 1996, §6 (Magazine), at 13 (letter to the editor) (suggesting that "after some preliminary thinning of the field, where real difference could be observed, the final selection [for elite colleges should be made by drawing names from a hat. This would be faster, fairer, and most of all, less damaging to the self esteem of those not admitted.").

of selection that is integrated into the day-to-day functioning of the organization.²⁵⁴ Recent developments in the assessment area, such as portfolio-based and authentic assessment, move in this direction.²⁵⁵ These might build on the tradition and virtues of apprenticeship, and indeed might "more closely resemble traditional apprenticeship measures than formal testing."²⁵⁶ They would build from and acknowledge the effects of context on performance and the importance of measuring performance in relation to context.

To take the next step in developing an experience-based approach to opportunity and assessment, it would be necessary to consider the needs, interests, and possibilities of the particular institutional setting. The task would be to figure out how to create opportunities to do work over an extended period of time that would provide a basis for assessing performance in that work. At an organizational level, this approach also integrates the process of defining standards and selecting applicants into the day-to-day work of the organization. Those who must work with or be served by the individuals selected would have more to say about what should be done, and whether particular individuals perform those functions well. In this sense, the model we describe is more participatory and democratic than the traditional top-down model of using tests to rank individuals for selection.²⁵⁷

A critical component to this selection process is the capacity to develop fair, reliable, and accountable mechanisms for assessing performance in these positions. We are not suggesting that the current approach of relying on supervisors' individual performance assessments would

254. See GARDNER, *supra* note 21, at 174-75 ("Rather than being imposed 'externally' at odd times during the year, assessment ought to become part of the natural learning environment."); *Applying Performance Assessment: The Work Sampling System*, FAIRTEST EXAMINER, Fall/Winter 1995-96, at 16 (describing promising uses of the work sampling system in elementary and secondary education).

255. See, e.g., GARDNER, *supra* note 21, at 171 ("Rather than attempting to devise more and better 'creativity tests,' researchers have instead begun to examine more closely what actually happens when individuals are engaged in problem-solving or problem-finding activities." (citation omitted)); MITCHELL, *supra* note 252, at 103-31 (describing portfolio-based assessment in education).

256. GARDNER, *supra* note 21, at 173.

257. See, e.g., Sabel, *A Measure of Federalism*, *supra* note 228, at 30. Sabel argues:

The more decentralized decision making, the more important that goals and the indicators by which their achievement is measured be continuously revised to reflect learning through exercise of local autonomy. Assessment becomes the continual adjustment of performance measures in the light of experience as a way to redefine joint goals while simultaneously evaluating progress.

Id. at 5; see also Thomas & Ely, *supra* note 248, at 2-4.

Some might say that this approach is nothing more than traditional affirmative action. The major difference, however, is that these organizations usually move beyond a concern simply with numbers. They focus as well on issues of mentoring and career development. Diversity concerns are often integrated into ongoing training and development efforts. Furthermore, diversity is seen as an issue for all levels of the organization.

Id. at 2.

fulfill this function. On the contrary, such individualized, unstructured judgments often reflect bias and have been found to be unreliable measures of performance.²⁵⁸ The challenge posed by this move is to develop systems of accountable decision making that minimize the expression of bias, and structure judgment around identified, although not static, norms. For each assessment, decision makers would articulate criteria of successful performance, document activities and tasks relevant to the judgment, assess candidates in relation to those criteria, and offer sufficient information about the candidates' performance to enable others to exercise independent judgment.²⁵⁹

For this model to work, institutions would also need to change the relationship of race, gender, and other categories of exclusion to the overall decision-making process. Institutions would continue to assess the impact of various selection processes on traditionally excluded groups. However, institutions would use that information in different ways. Rather than operating as an add-on, after-the-fact response to failures of the overall process, race and gender would serve as both a signal of organizational failure and a catalyst of organizational innovation. By serving this signaling function, race and gender would assist institutions in pursuing inclusiveness for traditional beneficiaries of af-

258. See Virginia E. O'Leary & Randal D. Hansen, *Performance Evaluation: A Social-Psychological Perspective*, in *PERFORMANCE MEASUREMENT AND THEORY* 197, 198 (Frank Landy et al. eds., 1983) (citation omitted):

[A]bstract representations or schemas that an individual has for a group (based on race, sex, nationality, or in-group—out-group membership) not only guide our search for new information and direct our attention to specific behaviors but also affect our memory for events and our distribution of rewards—two outcomes central to the performance evaluation process.

259. See LINDA DARLING-HAMMOND ET AL., *A LICENSE TO TEACH: BUILDING A PROFESSION FOR 21ST-CENTURY SCHOOLS* 71-88 (1995) (describing emergence of new assessment methods in teaching that proceed from contextually determined definitions of successful performance and that combine on-the-job evaluation, portfolio-based assessment, simulations, interviews, and paper-and-pencil tests); Linda Darling-Hammond, *Performance-Based Assessment and Educational Equity*, 64 *HARV. EDUC. REV.* 5, 18-19 (1994) (describing importance of "teacher and student involvement in and control over assessment strategies and uses" and the importance of developing the capacity "to evaluate and eliminate sources of unfair bias in [the] development and scoring of instructionally embedded assessments").

In the employment area, this approach may be easier to imagine within the context of promotion. Employers could create opportunities for current employees interested in new positions to experiment with new tasks and roles, either as part of training or on an experimental basis. Entry-level positions would require some additional creativity and thought. Employers could create training programs, work-study programs, or participate in field placements with schools and colleges to create work opportunities that would likely be part of the selection process.

In the education context, this approach has its most direct applicability in the classroom context: how do teachers evaluate the performance of students? In the context of admissions to institutions of higher learning, the model suggests decreasing reliance on standardized tests and increasing the emphasis on identifying the diverse range of skills, backgrounds, and experiences that would comprise a dynamic learning environment.

firmative action *and* in pursuing overall institutional fairness and productivity.

This approach of embedding concern for inclusion and diversity into the decision-making process would require identifying the key decision points of selection and then affording sufficient participation to assure that issues of inclusion and diversity are addressed at each level.²⁶⁰ The challenge would be to highlight and harness the synergy between inclusiveness and productivity.

Before we justify this proposal in relation to the principles of merit and fairness, it may be helpful to provide some illustrations of selection practices that resemble the model we set forth. These stories may help readers visualize what this process might look like, and prompt brainstorming about other creative examples of integrated and embedded approaches to selection.

What would an integrated approach look like? The story of one woman—call her Bernice—who is now the general counsel of a major financial institution, illustrates this approach.²⁶¹ Initially, she was hired as local general counsel to a bank. She was hired from her position as partner in a prestigious law firm, where she had reached the glass ceiling. In both rainmaking and firm power she perceived impenetrable barriers.

Bernice ultimately became general counsel to a major national corporation that previously had no women in high-level management positions. Her promotion resulted from the opportunities presented in an extremely interactive and extended selection process. Her local bank merged with a larger company. In part to create the appearance of including women, she was permitted to compete for the job of overall general counsel. Three lawyers shared the position for nine months. She did not view herself as in the running for the final cut.

During this time period, Bernice had a series of contacts with high-level corporate officials, contacts she never would have had without this probationary team approach. As it turned out, Bernice was able to deal unusually well with a series of crises. If standard criteria had been used to select, such as recommendations and interpersonal contacts, it is

260. In the employment context, these decision points could include (1) identifying a recruitment strategy, (2) identifying performance goals and criteria, and (3) designing and allocating training opportunities. Institutions would undertake to design decision-making groups that are themselves inclusive and diverse in their composition. The institutional caretakers of inclusiveness and diversity, such as affirmative action officers, human resource managers, or in-house counsel, could participate in helping design processes of recruitment and selection that minimize bias and hold decision makers accountable. This process could include seeking the input and involvement of the communities from which institutions seek to draw prospective students; employees could also be included in the process of recruiting and selecting applicants. Similarly, the increasingly diverse clientele to be served by institutions could play a role in shaping the function.

261. The actual identities of the individuals described below have been changed.

doubtful she would have been picked. However, teamwork, decentralized management, and collaborative and flexible working relationships allowed her to develop the contacts and experiences that trained her. The opportunity to interact over a period of time allowed her to demonstrate her strengths to those who made promotion decisions. Bernice did not know she had those strengths until she took the job.²⁶²

Now, as general counsel, she is positioned to influence profoundly opportunities for women, and corporate culture in general.²⁶³ She can structure the same kind of collaborative decision making in selection that provided her the opportunity to work her way into the job. She determines who is promoted within the legal department, and who is hired as outside counsel. She is also in a position to influence how women are assessed as managers within the company.

This story illustrates the potential of integrating concerns about diversity into the process of recruitment and selection. It also shows the value of using performance to assess performance. The second story of structured informal decision making in selection comes from the *Wall Street Journal*.²⁶⁴ It is the story of a black man who built a multimillion dollar business from a bicycle and an acute sense of how to spot, train, and continually reinforce diverse diamonds in the rough.

Lewis Roland quit his job as an academic administrator to start a same-day delivery service business. His business training consisted of watching his father manage the family pool hall and restaurant and supervising a college prep program called Upward Bound in an inner-city section of Newark. He trained local guidance counselors to spot "the student who was bright as hell, but who had no self-concept."²⁶⁵ In a business with a very tight labor market, he parlayed this capacity, coupled with a willingness to look in places others had overlooked, into a major competitive advantage.

Roland has developed a remarkably participatory and performance-based system of selection. He trains employees as scouts for new employees. He also relies heavily on community leaders to identify and refer individuals "with positive attitudes" and ability.²⁶⁶ Once they are identified, he sets up an interview process that replicates many of the aspects and challenges of the job of delivery. A "candidate is in-

262. She learned that she was proficient in skills that she did not previously identify as lawyering skills: problem solving, thinking about the public-relations management of crises, strategic planning, and dealing with internal disruption stemming from crisis and change.

263. See Grace M. Giesel, *The Business Client Is a Woman: The Effect of Women as In-House Counsel on Women in Law Firms and the Legal Profession*, 72 NEB. L. REV. 760 (1993) (discussing how the role of women as general counsel will improve the status of women in the legal profession).

264. Thomas Petzinger Jr., *Lewis Roland's Knack for Finding Truckers Keeps Firm Rolling*, WALL ST. J., Dec. 8, 1995, at B1.

265. *Id.*

266. *Id.*

structed to phone at a precise hour. The failure to do so means that the applicant could never be trusted with a 'just-in-time' delivery. If a candidate is invited to the office, Dr. Roland may keep him waiting in order to observe how the candidate handles it."²⁶⁷ This on-the-job interview is then replicated with constant meetings that rehearse and reinforce the values crucial to successful performance.

Roland institutionalizes both the independent value and the economic payoff of non-traditional selection practices and racial inclusion.²⁶⁸ His selection plan devotes resources up front to finding people on whom others gave up. Roland then creates circumstances to provide applicants the opportunity to show their capabilities, providing him with the information needed to assess whether they can do the job. By involving the community in the process of selection, and engaging employees in the recruitment and selection process, Roland obtains information unavailable from conventional sources and builds into the selection process a responsiveness to both external and internal needs.

Our last example of the innovative potential of collaboration in selection and recruitment is still in progress. Lowell High School in San Francisco is a magnet public school that boasts distinguished alumni, including Supreme Court Justice Stephen Breyer.²⁶⁹ As a result of a court proceeding to desegregate the San Francisco public schools, admission to the school is supervised by a court consent decree. No one ethnic group can comprise more than forty percent of the population of any magnet school. Consequently, in 1993 admission to Lowell High School proceeded on a sliding scale. Chinese Americans were required to score sixty-six out of a total of sixty-nine to gain admittance; "other whites" and non-Chinese Asian Americans could qualify with a fifty-nine, blacks and Latinos with a fifty-six. As a result of pressure from the Chinese-American community, these cutoff scores and entry credentials were modified somewhat. But the school still employed race-based quotas to protect diversity.

A group of Chinese Americans challenged the consent decree; African Americans defended it. Both groups proceeded within a winner-take-all frame, which pits minority groups into competing factions.²⁷⁰

267. *Id.*

268. Roland is especially eager to find candidates from diverse backgrounds. He is not discouraged by traditional markers of disqualification, including criminal records, if he can identify in the potential applicant a strong motivation to succeed. *Id.*

269. See *supra* note 20 (describing Lowell High School).

270. The School District responded to the conflict by deciding to admit 70-80% of the entering freshman class solely on their grades and test scores; the remaining 20-30% would be chosen "by a selection committee on a 'value added' diversity basis, focusing on socio-economic status, middle school coursework, extra-curricular activities, and residency within the city." San Francisco Unified School District, Revision of Lowell High School Admission Process (Feb. 12, 1996) (transcript of school board resolution, on file with authors).

We asked our seminar students to buck the hypo. Could they escape the false choice of winner-take-all decision making? Several innovative solutions emerged. One student suggested the use of a lottery that would allow anyone with a score over fifty-six—the score below which students are not admitted—to compete for admission via random selection. If the school could demonstrate, however, that those with a perfect sixty-nine or close to it, or with some other quality likely to produce something the school values—such as achieve a seat on the U.S. Supreme Court, win recognition as a Westinghouse Science Finalist, or be admitted to a competitive college—then those names would be placed in the lottery twice or even three times.²⁷¹ Students emphasized the importance of having a diverse and engaged group of participants in the selection process, including students, alumni, teachers, community representatives, and parents.²⁷²

Other students suggested instituting a summer program for Lowell High School, which would then offer the opportunity for selection based on students' participation and performance over the summer. As discussed above, a similar preparatory program designed to enhance minority students' performance has been utilized with success at Georgia Tech.²⁷³ Also proposed was the idea of expanding the number of opportunities to participate in a school like Lowell by either upgrading other schools or instituting a system of rotating enrollment, with students each having a year at Lowell. Schools would be developed with differing emphases and strengths, so that the selection process would focus on matching students to the right environment, rather than permanently selecting "the best" based on a single standard. None of these alternatives is perfect, but we offer them for their potential to generate multiple ways of reexamining the school's admission policy, not just for Chinese Americans or blacks, but for everyone.²⁷⁴

Another possible strategy to implement a more experience-based approach in the educational context more generally would involve emphasizing transfers from community colleges as a significant part of

271. Cf. David M. Herszenhorn, *Boston School Shows Cities Another Way: Charter Concept Puts Teachers in Control*, N.Y. TIMES, Jan. 31, 1996, at B8 (reporting that judge overseeing admissions lottery picked students to fill new charter school).

272. Cf. *id.* (quoting Sarah Kass, principal of City on a Hill Charter school, as saying that schools should "forge links with the best of the city" and that "those closest to kids [should] determine what the mission of the school is").

273. For a discussion of this program, see *supra* note 240 and accompanying text.

274. What all of these approaches have in common is the recognition that the claims of both "merit" and "diversity" are legitimate. They do not proceed as "us" against "them." They do not assume that only one group wins, and that those who win, win all. They reject a zero-sum solution in favor of a "positive-sum" solution that accommodates more broadly the goals of diversity and genuine merit. No one needs to feel "entitled" to admission; nor should anyone feel unjustly excluded.

admissions.²⁷⁵ This would allow candidates with lower paper-and-pencil credentials to demonstrate through performance their ability to excel in the academic setting. More fundamentally, the admissions process at colleges could be transformed into a two-tier evaluation, with junior colleges serving as feeder schools for all four-year universities. Under such a system, no candidate would be chosen on the basis of test scores or paper applications, but rather would be required to earn a place at her chosen university through performance in lower-division classes. These approaches to selection capture our idea of a relationship integrating inclusion, selection, and productivity.

At the core of this integrative move is a functional theory of diversity animated both by principles of justice and fairness (the inclusion of marginalized groups and the minimization of bias) and by strategic concerns (improving productivity). It is crucial to this integration that decision makers and advocates understand and embrace a functional conception of diversity that builds on both normative and instrumental goals. In public discourse, diversity has become a catchall phrase or cliché used to substitute for a variety of goals, or a numerical concept that is equated with proportional representation.²⁷⁶ Too often, these two strands of diversity discussion remain separate, with those concerned about justice emphasizing racial and gender diversity as a project of remediation, and those concerned about productivity emphasizing differences in background and skills to the exclusion of race and gender inclusion. The absence of an articulated theory that links diversity to the goals of particular enterprises and to the project of racial justice has complicated public discussion and public policy making around race and gender issues. The next Section ties the model we propose to instrumental goals of improving productivity and normative goals of pursuing fair and inclusive institutions.

B. Integrating Selection and Productivity

Perhaps the most broadly persuasive instrumental argument for the approach we propose is that it has the potential to improve institutions' capacity to select productive workers, to pursue innovative and adaptive performance, and to respond to the demands of a changing economic

275. Cf. John Martinez, *The Use of Transfer Policies for Achieving Diversity in Law Schools*, 14 CHICANO-LATINO L. REV. 140 (1994) (describing the benefits of diversity-conscious law school transfer policies).

276. See generally Richard Delgado, *Affirmative Action as a Majoritarian Device: Or, Do You Really Want to Be a Role Model?*, 89 MICH. L. REV. 1222 (1991). For example, the court in *Hopwood v. Texas* rejected the concept of diversity as a basis for adopting an affirmative action plan. Its discussion lacked almost any reflection on the functional role diversity plays in higher education. It simply asserted that "[t]he use of race, in and of itself, to choose students simply achieves a student body that looks different." 78 F.3d 932, 945 (5th Cir. 1996), *cert. denied*, 116 S. Ct. 2582 (1996).

environment. The traditional approach treats selection as entirely outcome-oriented. Selection is an add-on cost necessary to obtain the human capital needed to function. Selection is not traditionally treated as an integral process that can otherwise contribute to productivity. It is certainly not considered as part of the strategic planning process or as a means of continual economic revitalization.²⁷⁷ Selection proceeds from the top down, and its success is measured by two criteria: the measurable quality of those selected and the cost of the selection process.

This purely outcome-driven lens on validity and efficiency takes an overly narrow and static view of selection. Selection processes that impose additional, up-front costs may in the long run be more efficient if, for example, those processes enhance an institution's capacity to select capable candidates. If more dynamic notions of efficiency are taken into account, selection processes that force institutions to internalize the costs of their exclusionary practices can also be justified.²⁷⁸ The conventional top-down approach short-circuits the capacity of selection to serve as a mechanism for feedback about an institution's performance and its need to adapt to changing conditions. It also impedes institutions from developing more responsive, integrated, and dynamically efficient selection processes.

At the very least, the move to performance-based selection would reduce reliance on standardized tests, which in most cases do not enable institutions to identify the most successful applicants. This approach would instead focus decision makers' attention on creating the scenarios and contexts necessary to make informed judgments about performance. This would improve the capacity of institutions to find people who are creative, adaptive, reliable, and committed, rather than just good at test-taking. In some instances, these structured opportunities could directly contribute to the productivity of the organization.

A more interactive process of selection also provides an ongoing opportunity to assess and monitor organizational performance and to perceive and react to the changing character and needs of clients and

277. See BRION, *supra* note 227, at 728 (finding that personnel departments typically do not participate in annual planning, workforce planning, or other aspects of organizational development).

278. Economists have begun to challenge the notion that individualistic, short-term measures adequately reflect a plausible theory of human cognition or a comprehensive theory of productivity. Cf. AMARTYA SEN, ON ETHICS AND ECONOMICS (1987) (arguing that economics fails to account for the influence of ethics on rational individual behavior); Amartya K. Sen, *Rational Fools: A Critique of the Behavioral Foundations of Economic Theory*, in BEYOND SELF-INTEREST 25, 34-36 (Jane J. Mansbridge ed., 1990); Myra H. Strober, *Can Feminist Thought Improve Economics? Rethinking Economics Through a Feminist Lens*, AEA PAPERS & PROC., May 1994, at 143 (arguing in favor of feminist economics as a critical discipline); Strober, *supra* note 147 (same). Other scholars have noted the significance of groups and status in motivating and explaining economic behavior. See Richard H. McAdams, *Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination*, 108 HARV. L. REV. 1003, 1007-08 (1995).

employees. It provides information learned through the process of selection to the rest of the organization. In the process of redefining the standards for recruitment, the organization also redefines how those already in the institution should function. Selection operates at the boundaries of the organization. It necessarily exposes decision makers to the environment they operate in, provides access to information about the world in which the organization operates, and forces choices about its relationship with that environment. The process of defining the standards for positions also reflects and reinscribes the organization's priorities and direction. Emphasizing one set of skills over another in the selection process communicates to employees and students how the organization defines good work. Thus, the selection process is one that provides the opportunity and challenge of continually redefining standards in relation to the stakeholders of the organization, both inside and outside.

The approach we are proposing may, and most likely will, cost more to implement in the short run.²⁷⁹ This more interactive and engaged method of decision making would certainly take more time and resources than the process of administering standardized tests. Indeed, perhaps the clearest virtue of standardized testing is its administrative convenience. Standardized testing is more efficient if efficiency is measured only in the short run and in relation to the cost of the enterprise. However, this narrow and static definition of efficiency is short-sighted and counterproductive. An investment of resources up front has the potential to enhance the overall productivity of the organization, both by identifying more productive individuals and by enabling the institution to adapt better to its changing environment. If we expand our focus beyond a single institution to society at large, the long-term efficiency of the approach we suggest becomes even more compelling. By expanding access to education and opportunity in the short run, we save considerable resources in the long run by avoiding the high costs of poverty, crime, and poor health associated with chronic unemployment.²⁸⁰

279. Cf. GARDNER, *supra* note 21, at 180 (responding to the claim that performance-based assessment is too costly, and challenging the idea that formal testing is cost-effective in the long run). Gardner adopts the estimate that "a move toward more qualitatively oriented forms of education (and perhaps also to higher-quality education) might increase costs by 10 to 15 percent but probably not more." *Id.* at 180.

280. See, e.g., Arenson, *supra* note 159, at A1:

[D]uring one year in the 1980's[sic], graduates admitted under open admissions earned a total of almost \$67 million more than they would have if the university's program had not been instituted. Their additional lifetime earnings were estimated at about \$2 billion.

"The benefits [in terms of enhanced income, taxpaying capacity, and quality of life] of the policy flowed to a broad cross section of New York City's ethnic communities, both minority and white." . . .

Just as it is difficult to locate one standardized measure that accurately predicts performance, it will be difficult to isolate a single measure to demonstrate the greater long-term productivity of a more integrated approach to selection.²⁸¹ The best we can do at this point is to highlight the evidence of heightened productivity in companies that have embraced more participatory and interactive decision making.²⁸² We can also rely on the intuitive and theoretical power of an approach that emphasizes functional capacity as indicated by actual performance, rather than a numerical ranking derived from arbitrary and discriminatory criteria.

We believe that it is crucial to rethink merit and fairness in ways that transform institutions, rather than simply justify targeted and visible departures for race and gender. The current political campaign to delegitimize affirmative action programs as "racial preferences," although not the motivation for this move, certainly underscores its timeliness. We propose to internalize the critique of merit to the overall system of selection. We seek to develop not another test, but a new approach to selection that rethinks meritocracy as it is currently defined, and that casts race, gender, and other categories of exclusion in the role of signaling the direction for more inclusive, fair, and functional institutions.

C. A Functional Theory of Diversity

By embracing and harnessing difference, institutions can increase their functional capacity and their capacity to achieve genuine inclusion. Underlying this claim is an emerging functional theory of diversity, which rests on several assumptions. First, many, although not all, members of groups that have been formally excluded share experiences

281. Cf. Sabel, *A Measure of Federalism*, *supra* note 228, at 11-12 (discussing the difficulty of assessing the performance of training centers designed to improve business productivity).

282. Recent studies find that:

the adoption of a coherent system of these new work practices, including work teams, flexible job assignments, employment security, training in multiple jobs, and extensive reliance on incentive pay, produces substantially higher levels of productivity than do more "traditional" approaches involving narrow job definitions, strict work rules, and hourly pay with close supervision.

CASEY ICHNIEWSKI ET AL., ABSTRACT, THE EFFECT OF HUMAN RESOURCE MANAGEMENT PRACTICES ON PRODUCTIVITY (National Bureau of Economic Research Working Paper No. 5333, 1995); see also Edward P. Lazear, *Compensation, Productivity, and the New Economics of Personnel*, in RESEARCH FRONTIERS IN INDUSTRIAL RELATIONS AND HUMAN RESOURCES 341 (David Lewin et al. eds., 1992); Ben Craig & John Pencavel, *Participation and Productivity: A Comparison of Worker Cooperatives and Conventional Firms in the Plywood Industry*, in BROOKINGS PAPERS ON ECONOMIC ACTIVITY: MICROECONOMICS 121, 158 (1995) (finding that in the plywood industry cooperatives are more efficient than the principal conventional firms by 6% to 14%); MICHAEL DERTOUZOS ET AL., MIT COMMISSION ON INDUSTRIAL PRODUCTIVITY, MADE IN AMERICA: REGAINING THE PRODUCTIVE EDGE 111, 140 (1989) (arguing that cooperation of underdeveloped countries is a major obstacle to technological innovation and the improvement of industrial performance).

and cultures that influence their perception and mode of interacting in the world. Many people of color continue to express shared interests and perspectives.²⁸³ Second, at least in some instances, there is a community that articulates and identifies with a common set of concerns and interests. There are patterns of understanding, culture, and behavior that correlate with group membership and form the basis for community.²⁸⁴ Third, introducing these diverse viewpoints or perspectives can break down barriers to understanding, especially if there are opportunities for repeat encounters. Familiarity can breed respect.²⁸⁵ The benefits of diversity can only be reaped if we have processes of decision making that enable a range of views and perspectives to influence the decision-making process and outcome over time.²⁸⁶ Trust is necessary for constructive problem-solving; trust can only develop where conditions permit individuation, opportunity for correction, and genuine communication.²⁸⁷ Even for those committed to a progressive, universal set of reforms, deliberative diversity—if it encourages conditions for mutual trust—is crucial.²⁸⁸ Fourth, systemic problems may only become

283. One place where this group identification is both visible and quantifiable is in the widely documented preferences that blacks as a group express for black candidates. Samuel Issacharoff, *Polarized Voting and the Political Process: The Transformation of Voting Rights Jurisprudence*, 90 MICH. L. REV. 1833 (1992) (documenting the phenomenon of racial bloc voting). Bloc voting patterns are pervasive and demonstrate that many people of color, when given a choice, prefer to be represented by others who look like them and/or share their experience of discrimination. See Guinier, *The Triumph of Tokenism*, *supra* note 217. Social science studies reveal a comfort factor that extends beyond just racial minority groups: people feel more relaxed with those who seem familiar. See Guinier et al., *supra* note 68, at 74.

284. Lani Guinier, *[E]racing Democracy: The Voting Rights Cases*, 108 Harv. L. Rev. 109, 129-30 (1994) ("[M]inority group representation is not purely cultural, historical, or biological; it also has a political component. Group members may identify collectively along a common axis and organize to promote common interests in ways similar to other political associations."); see also CAROL M. SWAIN, *BLACK FACES, BLACK INTERESTS* 6-7 (1993).

285. See Guinier, *The Triumph of Tokenism*, *supra* note 217, at 1114 & nn.174-75 (noting that white constituents grow more willing to vote for incumbent black candidates once they become familiar with their performance and begin to see their competence at the things the constituents value).

286. See ANNE PHILLIPS, *THE POLITICS OF PRESENCE* 145-91 (1995). Phillips argues that a deliberative and representative process of inclusion does not guarantee that the concerns of people of color or women will always be considered. But such a process at least provides the possibility that representativeness can become a mechanism for channelling the concerns of those previously excluded. While diversity goals should not assume that all members of marginalized groups think alike, they should recognize that many of them share common perspectives. *Id.*

287. See Michael Walzer, *Multiculturalism and Individualism*, DISSENT, Spring 1994, at 185, 191 (arguing for importance of preserving both individual and group associations); see also Lani Guinier, *More Democracy*, 1995 U. CHI. LEGAL F. 1, 16-22 & n.13 (noting "socialization" effect of deliberation).

288. See, e.g., Thomas B. Edsall, *Public Grows More Receptive to Anti-Government Message*, WASH. POST, Jan. 31, 1996, at A1 (citing survey showing that rising levels of distrust among all Americans is particularly damaging to the traditional Democratic coalition because "Democrats need to build coalitions crossing racial and ethnic boundaries, coalitions for which trust [is] a crucial ingredient").

visible through the lens of race and gender. Finally, diverse approaches to problem solving can create innovative, creative solutions that otherwise would not emerge.

Functional diversity, if embraced as a central component of decision making, can create a synergy between the instrumental values of innovation and productivity, and the normative values of inclusion and fairness.

1. *Diversity, Innovation, and Productivity*

Research suggests that diversity is an independent value in generating creative solutions to problems. This aspect of diversity focuses on the interactive dynamic among individuals with different vantage points, skills, or values. Studies have shown that work-team heterogeneity promotes more critical strategic analysis, creativity, innovation, and high-quality decisions.²⁸⁹ Recent mathematical work suggests that participation of groups with different prior beliefs or predispositions in decision making improves the quality of the decision for everyone.²⁹⁰ Available research on jury deliberations supports the contention that diversity of participants contributes to improved deliberation.²⁹¹ A jury comprised of people from diverse backgrounds has more accurate recall and "more nuanced understanding of the behavior of the parties than [a more homogeneous jury]."²⁹²

Diversity in culture, style, and background also enhances the knowledge base and repertoire of skills and responses available to a particular group or institution.²⁹³ Indeed, the critique of the testocracy highlights the importance of reorienting institutions to make full use of the variety of ways to perform particular functions, some of which are undervalued by one-size-fits-all approaches to selection. Including

289. L. Richard Hoffman & Norman R.F. Maier, *Quality and Acceptance of Problem Solutions by Members of Homogeneous and Heterogeneous Groups*, 62 J. ABNORMAL & SOC. PSYCHOL. 401 (1961); see also Karen A. Bantel & Susan E. Jackson, *Top Management and Innovations in Banking: Does the Composition of the Top Team Make a Difference?*, 10 STRATEGIC MGMT. J. 107 (1989) (finding that diversity of functional backgrounds of top management teams was associated with organizational innovation); L. Richard Hoffman, *Applying Experimental Research on Group Problem Solving to Organizations*, 15 J. APPLIED BEHAVIORAL SCI. 375 (1979); Sumita Raghuram & Raghu Garud, *The Vicious and Virtuous Facets of Workforce Diversity*, in SELECTED RESEARCH ON WORK TEAM DIVERSITY 155, 156 (Marian N. Ruderman et al. eds., 1995) (finding that heterogeneous teams bring multiple perspectives to tasks and thereby outperform homogeneous teams in generating ideas).

290. Michael Suk-Young Chwe, *Taking Turns and Majority Rule*, Guinier and Condorcet, Minority Rights and Majority Welfare (Dec. 1995) (unpublished manuscript, on file with authors).

291. Kelly Darr, *Jury Deliberations and Successful Multiracial Problem Solving* (Mar. 28, 1996) (unpublished manuscript, on file with authors).

292. Jonathan D. Casper, *Restructuring the Traditional Civil Jury: The Effects of Changes in Composition and Procedures*, in VERDICT: ASSESSING THE CIVIL JURY SYSTEM 420 (Robert E. Litan ed., 1993).

293. See Raghuram and Garud, *supra* note 289, at 156.

people with different tendencies, styles, and approaches enhances flexibility and expands the repertoire of skills and functions that an institution can effectively pursue. Diversity offers new ideas and approaches that can enhance institutions' capacity to perform and innovate.²⁹⁴ The example of the Los Angeles Police Department, discussed above, illustrates this theory.²⁹⁵ The benefits of racial and gender diversity may be most obvious in the educational and human services areas, where customers, clients, and perspectives may themselves be identified by race and gender.²⁹⁶

Racial and cultural diversity in a workforce can also provide opportunities for companies marketing products that serve racially and culturally diverse client groups. As David Thomas and Robin Ely have documented, customers and clients from different racial, ethnic, and cultural communities constitute distinctive market niches that companies have sought to address by diversifying their workforces.²⁹⁷

2. *Race and Gender as Signifiers*

The experience of those who have been excluded or marginalized often signals more general or systemic problems that affect a much larger group and may hurt the organization's overall productivity. Sometimes, these problems only become visible through the lens of race and gender. The affirmative action critique of existing merit standards is only one example of this signaling function that race and gender can play. The unfairness and invalidity of one-size-fits-all approaches may be most visible when applied to women and people of color, but this critique of traditional approaches to selection can open up opportunities

294. See NEIL L. RUDENSTINE, HARVARD UNIVERSITY, THE PRESIDENT'S REPORT: "DIVERSITY AND LEARNING" (1996) (arguing that diversity is the "substance from which much human learning, understanding, and wisdom derive. It offers one of the most powerful ways of creating the intellectual energy and robustness that lead to greater knowledge, as well as the tolerance and mutual respect that are so essential to the maintenance of our civic society.").

295. See *supra* notes 132-137 and accompanying text (documenting that women and people of color tend to use different, more interactive styles of policing that check the tendency toward brutality and facilitate community policing).

296. See RUDENSTINE, *supra* note 294:

The PhD student who becomes a teacher of science, art, or economics at an undergraduate college (no less than the general practitioner of medicine, or the inner-city minister) must be prepared to understand and work with many individuals, over decades, who will have a multiplicity of opinions, cultural perspectives, and convictions about life. . . . [T]he realities of our time require forms of education that are broad in their human dimensions, as well as powerful in their intellectual content.

297. Thomas & Ely, *supra* note 248; see also Michael Diamond, *Slow Climb: Big Companies Such As Procter & Gamble Have Worked Hard to Hire More Minorities*, CINCINNATI BUS. COURIER, Jan. 11, 1993, at 1 ("A diverse team produces diverse ideas. . . . We are a consumer products company, and consumers are diverse. It's a very bottom-line need.").

to working-class whites and others who also have been inappropriately marginalized.²⁹⁸

Race and gender complaints also serve as a signal of more general organizational dysfunction. High rates of discrimination complaints may be symptomatic of more general management problems, such as poor organization or arbitrary treatment of workers.²⁹⁹ Because race and gender complaints often serve as the only visible source of information about these patterns of unfairness and counterproductivity, these group categories of analysis remain necessary from an institutional perspective. The need to respond to the exclusion of identifiable groups can prompt institutional self-evaluation and change.

Women's experience in law school provides an example of the crucial signaling function race and gender categories play. The recent studies documenting the experience of many women in law school as silencing and exclusionary reveal patterns of problems that many men experience as well.³⁰⁰ Similarly, sexual harassment of graduate students sometimes reveals a more general institutional inadequacy that would otherwise remain hidden. Faculty and students frequently lack shared understandings about fair, respectful, non-exploitative supervisory relationships between graduate students and their faculty advisors. Addressing sexual harassment—a problem ordinarily associated with women—can prompt a conversation on ways to promote productive and successful working relationships more generally.³⁰¹ In this way, race and gender exclusion can serve as a window, enabling us to see more general inadequacies in our structures and processes of decision making.

Race and gender constitute visible markers of diversity. As such, they can serve as a continual check or constraint on decision makers' impulse to revert to one-size-fits-all approaches to selection.

298. For example, the practice of job posting and advertisement opened up opportunities that were previously filled through closed, old-boy networks. See *supra* note 218 and accompanying text.

299. Lauren B. Edelman et al., *Internal Dispute Resolution: The Transformation of Civil Rights in the Workplace*, 27 L. & Soc'y REV. 497, 515-19 (1993) (arguing that discrimination complaints in organizations are often interpreted as merely reflecting bad management generally).

300. See Susan P. Sturm, *From Gladiators to Problem Solvers: Women, the Academy and the Legal Profession*, DUKE J. GENDER L. & POL'Y (forthcoming 1996).

301. See SUSAN STURM ET AL., UNIVERSITY OF PENNSYLVANIA, REPORT OF THE WORKING GROUP ON IMPLEMENTATION OF THE SEXUAL HARASSMENT POLICY (1994). In one recent example drawn from consulting experience, a number of employees were being sexually harassed by a supervisor who was engaging in a range of unprofessional behavior that was interfering with workers' ability to perform. By raising the issue of sexual harassment as a group and placing it in the context of respectful working relationships, the employees were able to stop the harassment and create a new set of expectations about day-to-day interactions.

3. *The Continued Significance of Underrepresented Groups*

Our commitment to using the lens of the margins to rethink the whole can be questioned from the perspective of economic revitalization and economic empowerment. Why does group status and performance continue to be crucial in assessing the adequacy of selection criteria? If we are successful in transforming the discourse and practice of merit and selection for everyone, why are race, gender, and other categories of exclusion still relevant to the discussion? If we continue to rely on such categories as race and gender, how do we avoid essentializing members of these groups and their performance?

In responding to this question, we take the world as it currently exists. Women and people of color have long been excluded and marginalized, and continue to experience exclusion in many institutional settings.³⁰² Race continues to be a divisive issue for many Americans, one that prompts skepticism and mistrust. The workforce is becoming increasingly diverse: almost two-thirds of entrants to the civilian workforce in the period between 1992 and 2005 are projected to be women and racial minorities.³⁰³ Our continued focus on race and gender moves forward from the current legal and organizational landscape. In many institutions, particularly those that are private and non-union, categories such as race and gender offer the only avenue for challenging decisions and practices.

Under these conditions, race- and gender-based inquiries continue to form the cornerstone of an integrated approach to a progressive economic agenda. Many members of marginalized groups predicate their willingness to participate in collaborative conversation on the majority's recognition of the ongoing significance of group-based exclusion. For members of historically excluded groups, a meaningful program of inclusion is a prerequisite to participating in ventures that benefit the whole community. Affirmative action has become a symbol of society's recognition of its responsibility for its history of legal disenfranchisement, and of the equal citizenship and respect of those who have historically been excluded.³⁰⁴ History shapes the perception and experience of those who have experienced formal exclusion, and this historic pattern of racial inequality will continue to be experienced unless it is affirmatively acknowledged and altered.

302. *E.g.*, HACKER, *supra* note 48, at 109-13 (documenting continued exclusion and underrepresentation of blacks).

303. COMMISSION ON THE FUTURE OF WORKER-MANAGEMENT RELATIONS, *supra* note 223, at 12.

304. Patricia Williams makes this point eloquently in *THE ALCHEMY OF RACE AND RIGHTS* (1991).

Without the cooperation of those concerned with race and gender justice in building this new progressive agenda, the dialogue will continue to be polarized, divisive, and adversarial. Unless we can build the concerns of racial and gender inclusion into the process of collaboration, these issues will continue to be addressed in adversarial settings that undermine the capacity of institutions to adapt to changing conditions.

In addition, research consistently shows that ignoring patterns of racial and gender exclusion causes these patterns to recur and dominate. A proven method of minimizing the expression of bias in decision making consists of reminding decision makers of the risk of bias or exclusion and requiring them to engage in fair, unbiased decision making.³⁰⁵ Unless we continue to pay attention to the impact of our decisions on members of groups that are the target of subtle bias and exclusion, those group members will continue to be marginalized.

Finally, we do not propose a categorical approach that is timeless and universal.³⁰⁶ The role diversity plays in the functioning of particular institutions will vary, depending on the institution's mission, demographics, and history. Diversity, like race, is not a static, fixed concept, but rather one that takes on meaning in the context of particular circumstances and projects.³⁰⁷ We acknowledge that race, gender, and other categories of identity matter in some contexts and not in others.

Group characteristics do matter for many, although not all, group members, at least in the current world as we know it. As conditions of inclusion and exclusion change, so may the capacity and need for race and gender categories to serve as signals of the need for change. At the

305. See Konrad & Linnehan, *supra* note 203, at 795 (summarizing research showing that evaluators may show less bias when identity-conscious structures are in place); see also *id.* at 807 ("Identity-conscious structures, but not identity blind structures, were positively associated with indicators of the employment status of women and people of color in the organizations studied. . . . [I]dentity-conscious structures are needed to ameliorate the biases of decision makers and reward systems . . ."); Madeline E. Heilman et al., *Has Anything Changed? Current Characterizations of Men, Women, and Managers*, 74 J. APPLIED PSYCHOL. 935, 936 (1989) ("[U]nder conditions in which performance effectiveness is verified, women are not depicted in traditionally stereotypic terms.").

306. Indeed, our approach, while it affords group membership a presumption of representativeness or treats it as a precondition for deliberative diversity, views this presumption as rebuttable. In other words, those group members who do not perform the functions of "representation" or mentoring could be rotated off a decision-making council or not hired as "role models"; conversely, non-group members who do perform these functions could establish their "diversity bona fides."

In tentatively suggesting a contextual and functional view of diversity, we understand the need to respect the integrity and autonomy of group members who occupy, or compete for, positions. We recognize the need not to create additional hurdles that once again only target people of color or women. Thus, these requirements of accountability should not be reserved only for group members, but should be built into the functions that every member of a decision-making council is expected to perform.

307. As others have noted, diversity is not a value in a vacuum. It must be "a mediating principle that informs us as to which differences matter." Foster, *supra* note 8, at 111.

very least, assuring women and people of color a place at the table increases the likelihood that the perspectives and experiences of these communities will influence the process and the result of decision making.

D. Selection and Substantive Fairness

We believe that by using the margins to rethink the whole, it becomes possible to achieve both the perception and the reality of fairness. Institutions that take seriously the challenge of using performance to develop opportunity can, we argue, be genuinely and substantively fair and inclusive.

First, this approach is less likely to exclude people who can actually perform in the position.³⁰⁸ This functional approach to selection reduces the importance of criteria that have excluded women and people of color and favored wealthier applicants. It enables previously excluded people to "show their stuff." It also normalizes a selection process that is fairer for women and people of color. Second, by questioning and rethinking the standards of selection for everyone, this approach destabilizes the idea that the existing meritocracy is itself fair. Embedding the role of diversity enables other people to see how benefiting women and people of color benefits them. Third, the functional approach has the potential to create a participatory and accountable selection process, which can enhance individuals' autonomy and institutions' legitimacy.³⁰⁹

Finally, conditions for sustained contact, genuine collaboration, and fair assessment provide outgroup members a meaningful opportunity to learn, perform, and succeed.³¹⁰ Studies of multi-racial teamwork suggest that the opportunity to work as relative co-equals in interdependent, co-operative teams may also reduce bias.³¹¹ Indeed, carefully structured,

308. See Carla Seaquist, *Pete Wilson's Gorgeous Mosaic*, N.Y. TIMES, Aug. 15, 1995, at A17, describing San Diego's nationally acclaimed affirmative action program, which never utilized quotas. Women and members of minority groups became police officers, firefighters, truck drivers, electricians, water-treatment plant operators, park maintenance workers, and lifeguards. *Id.* "The city achieved balance by rooting out considerations unrelated to job performance." *Id.*

309. See Barenberg, *supra* note 232, at 893-904 (summarizing the normative arguments for workplace participation); Joshua Cohen, *The Economic Basis of Deliberative Democracy*, 6 Soc. PHIL. & POL'Y 25 (1989).

310. See David A. Thomas & Clayton P. Alderfer, *The Influence of Race on Career Dynamics: Theory and Research on Minority Career Experiences*, in HANDBOOK OF CAREER THEORY 133 (Michael B. Arthur et al. eds., 1988); see also Clayton P. Alderfer, *An Intergroup Perspective on Group Dynamics*, in HANDBOOK OF ORGANIZATIONAL BEHAVIOR 190 (Jay W. Lorsch ed., 1987).

311. See Samuel L. Gaertner et al., *The Contact Hypothesis: The Role of a Common Ingroup Identity on Reducing Intergroup Bias*, 25 SMALL GROUP RES. 224, 226 (1994); Samuel L. Gaertner et al., *How Does Cooperation Reduce Intergroup Bias?*, 59 J. PERSONALITY & SOC. PSYCHOL. 692 (1990). Similar studies have been conducted in schools. See Elliot Aronson & Diane Bridgeman, *Jigsaw Groups and the Desegregated Classroom: In Pursuit of Common Goals*, 5 PERSONALITY &

accountable, and participatory work groups may replicate the conditions most likely to reduce bias and permit genuine participation by women and people of color.³¹² Selection processes emphasizing problem-solving and performance may reduce the likelihood that stereotype anxiety will interfere with performance, a dynamic that has been identified with tests measuring intellectual ability.³¹³

We acknowledge that these new, more interactive and informal forms of selection and management rely explicitly on discretion and subjectivity. Preconceptions and biases will likely affect evaluations of performance in ways that often exclude women and people of color. We also recognize that unstructured discretion exercised without accountability or participation by diverse decision makers will likely reproduce biased and exclusionary results. However, these biases have not been eliminated by formal selection practices and paper-and-pencil tests. More importantly, the model of formal fairness that is outcome-driven, rule-bound, and centralized will not reach many of the places where women and people of color seek to enter.³¹⁴ If the economy is moving in the direction of creating and restructuring work along more team-oriented, participatory lines,³¹⁵ approaches to selection and performance must evolve that permit women and people of color to participate fairly and to succeed in this changing environment. Otherwise, women and people of color will remain on the margins of the new economy. Moreover, as business entities become more fluid and rely more on subcontracting and temporary work,³¹⁶ we must devise new and more interactive strategies for inclusion and empowerment that embrace a workforce existing in the margins of traditional legal categories. The exercise of discretion cannot and should not be eliminated. Instead, discretionary decision making must become the subject and site of participation, accountability, and creative problem-solving.

No system of accountability will eliminate bias. What it can do is create a context for exposing and minimizing the expression of bias.

SOC. PSYCHOL. BULL. 438 (1979); David W. Johnson & Roger T. Johnson, *Effects of Cooperative, Competitive, and Individualistic Learning Experiences On Cross-Ethnic Interaction and Friendships*, 118 J. SOC. PSYCHOL. 47 (1982).

312. Cf. BLUEPRINT FOR GENDER EQUITY, *supra* note 128, *passim* (recommending reforms that include women in decision making, articulate clear standards, and establish mechanisms of accountability to achieve the goal of diversifying the police force).

313. See *supra* note 92.

314. See Elizabeth Bartholet, *Application of Title VII to Jobs in High Places*, 95 HARV. L. REV. 947, 967-78 (1982) (discussing courts' reluctance to scrutinize high-level employment decisions); Deborah L. Rhode, *Perspectives on Professional Women*, 40 STAN. L. REV. 1163, 1193-94 (1988) (noting courts' deference to employers' judgments).

315. See, e.g., *supra* notes 228-235 and accompanying text.

316. COMMISSION ON THE FUTURE OF WORKER-MANAGEMENT RELATIONS, *supra* note 223, at 21-22 (noting increase in "contingent work" resulting from employer needs for increased flexibility and cost savings).

The model we set forth is designed to force decision makers to become self-conscious about their priorities, incentives, and structures of decision making. It moves from the idea that standing still in a rapidly changing world is risky, and that long-term success requires self-conscious experimentation. It understands that tension and uncertainty are inevitable in a complex and dynamic world and can be a source of innovation as well as conflict.

We also do not suggest that the process of moving in this direction will be linear or smooth. Although they offer tremendous potential for creativity, diverse working groups also pose particular tensions and challenges. But these challenges are inevitable, given the increasing diversity of the future workforce, and they restrain the tendency to fall back on simple but inadequate solutions. Uncertainty, tension, and conflict cannot be avoided, either in responding to the challenge of racial and gender inclusion or in developing the capacity to adapt to changing economic and technological times.

E. A Democratic Imperative

Finally, there is a normative basis for rethinking the relationship of race, gender, and class to selection procedures for work and school. Access to work and education is rapidly becoming a fundamental attribute of modern citizenship at the turn of the century. Work provides an identity that is valued by others. Work organizes and shapes the citizen's sense of self. Work legitimates. Virtually every aspect of citizenship is channeled through participation in the workplace. For most people, medical care, pensions, and social insurance are linked to workplace participation. In these ways, work has become a proxy for citizenship.³¹⁷

Increasingly, the opportunity to work in a non-contingent, full-time position that provides these benefits of citizenship depends on access to higher education. People who are not educated do not get jobs, and thus cannot participate in the responsibilities and benefits of citizenship. Moreover, those without the benefits of higher education increasingly work in shifting, temporary, and task-centered jobs. Such individuals may fail to develop a sense of personal worth, institutional or communal loyalty, or positive agency, all attributes essential to functioning as citizens.³¹⁸

317. See Richard Sennett, *A Place in the World: Work and Community in the New Economy* 1 (May 1, 1996) (paper presented at the Bellagio Conference on the Humanities and the Social Sciences, on file with authors). "When Hegel declared that everyone needs a place in the world, he was thinking about citizenship; no one can bear to be stateless. But there are other ways to satisfy the need for a place in the world, through one's work, or through living in a community." *Id.*

318. *Id.* at 6-10 (arguing that work is the foundation of self and that productive experience is a source of self-worth).

In addition, voting—the process that has traditionally served to permit participation and influence public decision making—does not afford individuals the capacity to deliberate and exercise much influence over the conditions of day-to-day life.³¹⁹ Without the opportunity to participate in intermediate institutions, such as places of work and schools, many citizens have no sense that their voices are being heard.³²⁰

This basic connection among work, education, and citizenship suggests that the screening process for employment and education has become the modern-day equivalent of eighteenth- and nineteenth-century screening processes for voting. In the colonial period and the first decades of independence, the franchise was generally restricted by race and gender to landed white males who owned property. In the late nineteenth century, voting was also conditioned on the capacity to pay and the ability to read.

Throughout our history, many jurists have argued that the poll tax and literacy requirements served legitimate interests by assuring that those who performed the fundamental and crucial responsibility of voting were capable and committed citizens.³²¹ As the second Justice Harlan observed, the poll tax was a valuable way to promote

civic responsibility, weeding out those who do not care enough about public affairs to pay \$1.50 or thereabouts a year for the exercise of the franchise. It is also arguable, indeed it was

319. See, e.g., Guinier, *supra* note 287; see also Lani Guinier, *Democracy as Theater*, COLUM. JOURNALISM REV., March 1996.

320. This is a complex argument that requires more elaboration than the limits of this article permit. Suffice it to state the obvious: we are experiencing a massive retreat from public life on many levels. The declining percentage of Americans who actually cast a ballot is just one measure. See David Glass et al., *Voter Turnout: An International Comparison*, PUB. OPINION, Dec./Jan. 1984, at 49, 49 ("Everyone knows that Americans vote less than citizens of other democratic countries."); FRANCES F. PIVEN & RICHARD A. CLOWARD, *WHY AMERICANS DON'T VOTE 4* (1988) ("The universe of actual voters in the United States is shrunken and skewed compared with the universe of formally enfranchised citizens. Only a little more than half of the eligible population votes in presidential elections, and fewer still vote in off-year elections."). Those who do vote tend to be better off and better educated. Non-voters are poorer and less well educated. "In sum, the active American electorate overrepresents those who have more, and underrepresents those who have less." *Id.* See also Richard L. Berke, *Nonvoters Are No More Alienated Than Voters*, *A Survey Shows*, N.Y. TIMES, May 30, 1996, at A21 (reporting that League of Women Voters survey found that voters and nonvoters are equally mistrustful of government, but that nonvoters are far less inclined to believe that their participation will make a difference to the outcome of an election).

If people are not voting, then our ability to forge democratic consensus suffers. If we are to retain legitimacy as a genuine democracy, we must either change the election system to encourage greater participation or engage people in other citizenship opportunities. Participation in intermediate institutions such as the workplace offers one alternative democratic space. This alternative is particularly appealing if work itself has become an important source of public identity. It is also appropriate if the workplace is the major public meeting ground for people who otherwise live and play separately by race and ethnicity.

321. See e.g., *Harper v. Virginia Bd. of Elections*, 383 U.S. 663, 685 (1966) (Harlan, J., dissenting); *Lassiter v. Northhampton County Bd. of Elections*, 360 U.S. 45 (1959).

probably accepted as sound political theory by a large percentage of Americans through most of our history, that people with some property have a deeper stake in community affairs, and are consequently more responsible, more educated, more knowledgeable, more worthy of confidence, than those without means, and that the community and Nation would be better managed if the franchise were restricted to such citizens.³²²

The current testocracy arguably operates as a modern-day poll tax, constricting the opportunities for participation in contemporary forms of citizenship.³²³ Like literacy tests, which promote the intelligent use of the ballot, general ability tests screen those who "deserve" or are "worthy of confidence" to work. This analogy between contemporary wealth-correlated testing criteria and eighteenth- and nineteenth-century voting practices is tentative and undeveloped. Yet, over time, our view of distributing opportunities for education and work may evolve just as our notions of democracy did. Property and poll-tax qualifications were eventually considered "unjustified discrimination in determining who may participate in political affairs or in the selection of public officials," a discrimination which undermines "the legitimacy of representative government."³²⁴ Just as "[v]oter qualifications have no relation to wealth nor to paying or not paying this or any other tax,"³²⁵ eventually opportunities for modern-day citizenship may be democratically available without regard to "wealth" or to the payment of "this or any other tax." If wealth "is not germane to one's ability to participate intelligently in the electoral process,"³²⁶ then arguably access to work and education should not be channeled by a wealth-related credentializing process, especially one with severe race and gender consequences.

If, as we believe, work and education are becoming basic components of citizenship, screens or barriers to participation should be drawn in the least exclusive manner consistent with the institution's mission. Access and opportunity to participate is critical to equipping citizens to fulfill their responsibilities, to respecting their status and autonomy as individuals, and to legitimating society's decisions as reflecting the participation of the community. This argument depends upon a particular

322. *Harper*, 383 U.S. at 685. The *Lassiter* Court upheld literacy tests on very similar grounds. 360 U.S. at 52.

323. See *Harper*, 383 U.S. at 666 (concluding that a State violates the Equal Protection Clause of the Fourteenth Amendment "whenever it makes the affluence of the voter or payment of any fee an electoral standard").

324. *Kramer v. Union Free Sch. Dist.*, 395 U.S. 621, 621 (1969).

325. *Harper*, 383 U.S. at 666.

326. *Id.* at 668.

view of fairness as legitimacy.³²⁷ In this sense, fairness means democratic decision making or the idea that people who feel they have a voice in the decision-making process are more likely to accept the ultimate decision, even if it is different from the one they initially supported.³²⁸

In the voting context, race-, gender- and wealth-related restrictions were lifted "only after wide public debate" about "the very nature of the type of society in which Americans wished to live."³²⁹ These barriers were invalidated because they came to be seen as unduly burdening access to this fundamental aspect of citizenship. Courts also recognized that these burdens, through the exercise of selective discretion by local officials, fell disproportionately on disempowered groups such as African Americans.³³⁰ Likewise, we believe a national debate on the terms of participation in equivalent forms of citizenship is long overdue. Just as "history has seen a continuing expansion of the scope of the right of suffrage in this country,"³³¹ so we would argue that twenty-first-century democracy will depend on a commensurate expansion of the scope of access to higher education and opportunities for on-the-job training. Even if there are justifications for requirements relating to the capacity to exercise citizenship responsibilities effectively, these requirements must be drawn in the most narrow way possible because of the importance of assuring democratic access and legitimacy to the distribution of citizenship opportunities and responsibilities.

V

CONCLUSION

We have argued that what began as a potentially noble effort to bring objectivity to the selection, hiring, and promotion process has instead deteriorated into a preoccupation with the false promise of quantitative measurement. We are mired in a testocracy that, in the name of merit, abstracts data from individuals, quantifies those individuals based on numerical rankings, exaggerates its ability to predict those individu-

327. Fairness is used as a substantive expression of the legitimacy of the process used to define the existing selection process. Fairness as legitimacy reflects the importance of genuine participation and real consent.

328. Cf. Foster, *supra* note 8, at 112 (emphasizing the importance of "[i]ncluding individuals from [previously excluded] groups in the deliberative decision-making processes of society's institutions [to] allow full participation for those who have been historically excluded from deciding by what standards they will be judged and defining what perspectives and outlooks they will have").

329. *Harper*, 383 U.S. at 684 (Harlan, J., dissenting).

330. See, e.g., *U.S. v. Louisiana*, 225 F. Supp. 353, 355-56 (E.D. La. 1963) (finding that the interpretation test as a prerequisite for registration "has been the highest, best-guarded, most effective barrier to Negro voting in Louisiana," and that the test "has no rational relation to measuring the ability of an elector to read and write"), *aff'd*, 380 U.S. 145 (1965).

331. *Reynolds v. Sims*, 377 U.S. 533, 544 (1964).

als' future performance, and then disguises under the rubric of "qualifications" the selection of those who are more socio-economically privileged. Conventional selection methods fail because they give preferences to people based on socio-economic position, and allocate positions in ways that do not reflect functional capacity.

Many proponents of merit-based decision making disparagingly link affirmative action to number-counting and group-based preferences. They assert that genuine merit requires admitting or hiring the best-qualified *individual*, a determination made only by ranking along a set of numerical assessments. They equate numerical quantification with merit, and treat merit as if it were the weather reduced to a single day's temperature: an observable phenomenon measured by a thermometer. The irony is that advocates of these assessment techniques engage in a sleight-of-hand logic that values number-counting over individuality and privileges certain groups over others.

Some affirmative action proponents challenge individual decision making as too fraught with bias. Because subjective decision making cannot be trusted, these supporters of affirmative action also prefer outcome-oriented numerical assessment. In essence, one side uses numerical proxies for group decision making and calls it merit; the other uses numerical proxies for group decision making and calls it fairness.

We have argued that selection should be structured to enable individuals to show what they can do and to enable decision makers to make decisions based on an individual's capacity to perform. Unless we are prepared to move to a lottery system for allocating opportunity, we cannot fairly and democratically avoid individual assessment that takes into account functionally relevant differences, and provides individuals the opportunity to demonstrate, in context, what they are capable of doing.

Our approach incorporates an equally dynamic and particularized view of race. Although race often functions as a signal of class-based exclusion, we do not propose shifting focus to class rather than race. This approach would fail to respond to the particular dynamics of racism that operate regardless of social class. It also would fail to link concerns of inclusion with the critical task of reshaping institutional priorities in order to value and benefit from functional diversity. As such, this approach would preserve intact a system of selection that continues to marginalize those who enter the terrain on different terms.

Instead, we seek to open up a conversation about issues that many people treat as resolved. Our institutions do not currently function as fair and functional meritocracies. Only by rethinking our assumptions about the current system and future possibilities can we move toward the ideals that so many Americans share. This enterprise offers the possibility of bringing together many who are adversaries in the current

affirmative action debate but share an interest in forging fairer, more inclusive, and more democratic institutions. It reconnects affirmative action to the innovative ideal. In this way, affirmative action can reclaim the historic relationship between racial justice and the revitalization of institutions to benefit everyone.